

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2004, 23:34:58 ; Search time 158 Seconds

(without alignments)  
615.289 Million cell updates/sec

Title: US-10-090-185-9

Perfect score: 1388

Sequence: 1 RCLMEERSLQTAATAAQAQG.....LNTQLKIKVCIDKSDGVAA 271

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_23Sep04:\*  
2: geneseqp19808:\*  
3: geneseqp19908:\*  
4: geneseqp20008:\*  
5: geneseqp20028:\*  
6: geneseqp20038:\*  
7: geneseqp20038:\*  
8: geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1388	100.0	271 4	AAV72841 Mouse Sta
2	1388	100.0	770 2	AAV72082 Mouse Sta
3	1388	100.0	770 2	AAW03176 Mouse Sta
4	1377	99.2	720 5	AAE22055 Human Sta
5	1377	99.2	769 5	ABBS7164 Mouse Isc
6	1377	99.2	769 5	AAE22054 Human Sta
7	1377	99.2	769 5	AAE22056 Human pro
8	1377	99.2	770 2	AAE22995 Mouse Iiv
9	1377	99.2	770 2	AAV03768 Human STA
10	1377	99.2	770 2	AAE12377 N-termina
11	1377	99.2	770 2	AAE14652 Murine ST
12	1377	99.2	770 2	ABG69497 Human bal
13	1377	99.2	770 2	ABU10476 Mouse STA
14	1377	99.2	770 2	ADN04365 Antipsori
15	1377	99.2	770 2	AAE58442 Lung cancr
16	1374	99.0	770 2	ADDA4738 Rat Prote
17	1372	98.8	770 2	AAE22993 Human pla
18	1372	98.8	770 2	AAE19964 Human sly
19	1372	98.8	770 2	AAE15174 Human Sta
20	1372	98.8	770 2	ADDA4740 Human pro
21	1290	92.9	252 4	AAV72846 Mouse Sta
22	1212	87.3	236 4	AAV72847 Mouse Sta
23	1172	84.4	229 4	AAV72850 Mouse Sta
24	1168	84.1	229 4	AAV72863 Mouse Sta
25	1167	84.1	229 4	AAV72862 Mouse Sta

26	1157.5	83.4	228 4	AAV72861 Mouse Sta
27	1146	82.6	229 4	AAV72860 Mouse Sta
28	1142	82.3	223 4	AAV72854 Mouse Sta
29	1094	78.8	213 4	AAV72851 Mouse Sta
30	943	67.9	185 4	AAV72855 Mouse Sta
31	899	64.8	176 4	AAV72848 Mouse Sta
32	723	52.1	143 4	AAV72849 Mouse Sta
33	669	48.2	129 4	AAV72856 Mouse Sta
34	654	47.1	749 5	AAV72852 Mouse Sta
35	653	47.0	128 4	AAV72852 Mouse Sta
36	647	46.6	268 4	AAV72844 Mouse Sta
37	647	46.6	680 6	ABR59713 Human sly
38	647	46.6	712 2	AAV72079 Human Sta
39	647	46.6	712 2	AAW03170 Human STA
40	647	46.6	712 2	AAW62995 Human Sta
41	647	46.6	712 6	ABU04747 Human exp
42	647	46.6	712 6	ABU04735 Human exp
43	647	46.6	712 6	ABU04745 Human exp
44	647	46.6	712 6	ABU04743 Human exp
45	647	46.6	712 8	ADH57035 Truncated

## ALIGNMENTS

## RESULT 1

AAV72841  
ID AAV72841 standard; protein; 271 AA.

AC AAV72841;

DT 31-MAY-2001 (first entry)

DE Mouse Stat3 protein fragment #2 (107-377 amino acids).

KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription; cellular transformation; dysproliferative disease; cancer; proinlasis; therapy.

XX

OS Mus musculus.

XX Key Location/Qualifiers

FT Region 24..48  
/note="Stat3-c-Jun interaction region 1; corresponds to 130-154 position of Stat3 protein"

FT Region 236..252  
/note="Stat3-c-Jun interaction region 2; corresponds to 342-358 position of Stat3 protein"

FT WO200116605-A2.

PN 08-MAR-2001.

PD 30-AUG-2000; 2000WO-US023822.

PF 31-AUG-1999; 99US-00387418.

PR (UYRQ ) UNIV ROCKEFELLER.

PA Zhang X, Horvath C, Wzreszcynska MH, Darnell JB;

PI WPI; 2001-226705/23.

DR Identifying an agent for use in modulating the interaction between

PT transcription factor c-Jun and a Stat3 protein.

PS Claim 65; Page 67-68; 86pp; English.

XX The present sequence is mouse Stat3 protein fragment containing 107-377

CC amino acids of Stat3 protein. This Stat3 fragment showed strong binding

CC to c-Jun protein in the cell extract. The invention relates to methods

CC for identifying interacting regions of transcription factors and methods

CC for identifying agents which modulates the interaction between a

CC transcription factor such as c-Jun and a Stat protein such as Stat-1 and  
 CC Stat-3, useful for modulating gene transcription e.g., cellular  
 CC transformation. These identifying agents are used in the treatment of  
 CC dysproliferative diseases and also for treating cancer and psoriasis. A  
 CC Stat protein comprises the N-terminal domain, coiled-coil domain, DNA  
 CC binding domain, linker domain, SH2 domain and transactivation domain  
 CC  
 XX Sequence 271 AA;

Query Match 100.0%; Score 1388; DB 4; Length 271;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-116;  
 Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RCMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRRVQDLEQKMKVE 60  
 DB 1 RCMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRRVQDLEQKMKVE 60  
 QY 61 NLDDDPENFKTKSGQDMQDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 120  
 DB 61 NLDDDPENFKTKSGQDMQDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 120  
 QY 121 AMEYVQKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLOTRQOIKKLE 180  
 DB 121 AMEYVQKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLOTRQOIKKLE 180  
 QY 181 LOOKSVYKGPPIYQHRPMLERIVELFRNLKSAFVVERQPCMPHMDRPLVKTGVQFT 240  
 DB 181 LOOKSVYKGPPIYQHRPMLERIVELFRNLKSAFVVERQPCMPHMDRPLVKTGVQFT 240  
 QY 241 TKVRLVKFPELNYOLKIVCICDKSGDVAA 271  
 DB 241 TKVRLVKFPELNYOLKIVCICDKSGDVAA 271

# RESULT 2

AA072082  
 ID AAR72082 standard; protein; 770 AA.

XX AAR72082;

AC 25-MAR-2003 (revised)  
 DT 27-SEP-1995 (first entry)

XX Mouse Stat3 (1996f).

DE Signal transducer and activator of transcription; STAT; 1996f; Stat3;  
 KW receptor recognition factor; transcription factor; cellular debilitation;  
 KW derangement; dysfunction; interferon-gamma.

XX Mus sp.

OS W09508629-A1.

XX 30-MAR-1995.

PD 26-SEP-1994; 94WO-US010849.

PF 24-SEP-1993; 93US-00126588.

PR 24-SEP-1993; 93US-00126595.

PR 11-MAR-1994; 94US-00212184.

XX (UYRQ) UNTIV ROCKEFELLER.

PI Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;

DR WPI; 1995-139598/18.

XX N-PSDB; AA089340.

PT Receptor recognition factor implicated in transcriptional stimulation of  
 PT genes - useful in drug screening assays and/or for treating cellular  
 PT debilitations, derangements and/or dysfunctions, etc.

PS Claim 1; Page 107-110; 160pp; English.

XX A fragment encoding the human Stat91 protein was used to screen a murine  
 CC thymus and spleen cDNA for homologous proteins. A highly homologous gene  
 CC (given in AA089338) was isolated that encoded a 91 kDa protein (AAR72080)  
 CC (Stat1) that was responsive to interferon-gamma. Using a fragment of the  
 CC mouse gene as probe, 2 additional members of the 113-91 family of  
 CC receptor recognition factor proteins were isolated. The 2 genes (AA089339  
 CC -40) were cloned in plasmids 136f1 and 198f6 and encoded proteins termed  
 CC Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-  
 CC 2003 to correct PN field.)

XX Sequence 770 AA;

Query Match 100.0%; Score 1388; DB 2; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 1e-115;  
 Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RCMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRRVQDLEQKMKVE 60  
 DB 107 RCMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRRVQDLEQKMKVE 166  
 QY 61 NLDDDPENFKTKSGQDMQDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 120  
 DB 167 NLDDDPENFKTKSGQDMQDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 226  
 QY 121 AMEYVQKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLOTRQOIKKLE 180  
 DB 227 AMEYVQKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLOTRQOIKKLE 286  
 QY 181 LOOKSVYKGPPIYQHRPMLERIVELFRNLKSAFVVERQPCMPHMDRPLVKTGVQFT 240  
 DB 287 LOOKSVYKGPPIYQHRPMLERIVELFRNLKSAFVVERQPCMPHMDRPLVKTGVQFT 346  
 QY 241 TKVRLVKFPELNYOLKIVCICDKSGDVAA 271  
 DB 347 TKVRLVKFPELNYOLKIVCICDKSGDVAA 377

# RESULT 3

AA003176  
 ID AAM03176 standard; protein; 770 AA.

XX AAM03176;

AC 24-OCT-1996 (first entry)

DE Mouse STAT4.

XX STAT; STAT4; signal transducer and activator of transcription;  
 KW DNA binding protein; ligand; receptor; oncogenesis; inflammation;  
 KW autoimmune disease; antagonist; therapy.

XX Mus sp.

OS W09620954-A2.

XX 11-JUL-1996.

PD 28-DEC-1995; 95WO-US017025.

PF 06-JAN-1995; 95US-00369796.

XX (UYRQ) UNTIV ROCKEFELLER.

PI Darnell JE, Wen Z, Horvath CM, Zhong Z;

DR WPI; 1996-333941/33.

XX

DR N-PSDB; AAT31280.  
 XX New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,  
 XX preventing or treating cellular dysfunction, e.g. oncogenesis,  
 PT inflammation, parasitic disease or autoimmunity.  
 XX  
 XX  
 PS Disclosure; Page 87-90; 138pp; English.  
 XX  
 XX Mouse signal transducer and activator of transcription (STAT) protein  
 CC STAT4 (AA03176) serves a dual purpose, i.e. signal transduction from  
 CC ligand-activated receptor kinase complexes followed by nuclear  
 CC translocation and DNA binding to activate transcription. Recombinant  
 CC STAT4 can be obtained using cDNA clone 198f6 (AAT31278) obtained from  
 CC splenic/thymic cells. STAT4 includes a DNA-binding domain (see also  
 CC AA03167) capable of both receptor recognition and message delivery via  
 CC DNA binding in a receptor-ligand specific manner. STAT proteins and their  
 CC binding domains (see also AA03165-75) are useful for screening  
 CC antagonists used to inhibit STAT-mediated signal transduction and  
 CC activation of transcription  
 CC  
 SQ Sequence 770 AA;  
 Query Match 100.0%; Score 1388; DB 2; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 1e-115;  
 Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RCLMESRLDTAATTAAGGGAHPTAAVTEKQOMLEOHLDVREKVDLEQKAYVE 60  
 DB 107 RCLMESRLDTAATTAAGGGAHPTAAVTEKQOMLEOHLDVREKVDLEQKAYVE 166  
 QY 61 NLQDDFDNRYKTKSGQGMODLNGNSVTRQKMOLEQMTALDQMRSTVSELAGLLS 120  
 DB 167 NLQDDFDNRYKTKSGQGMODLNGNSVTRQKMOLEQMTALDQMRSTVSELAGLLS 226  
 QY 121 AMEVYQKTLTBEALADMKRRPPIACIGPPNICDLRLNMTTSLAESQLQTRQOIKKLE 180  
 DB 227 AMEVYQKTLTBEALADMKRRPPIACIGPPNICDLRLNMTTSLAESQLQTRQOIKKLE 286  
 QY 181 LQOKSYKGDPIVQHRPMLBEIYELFRLMKSAFVEROPCMHPRPPIVYIKTVQFT 240  
 DB 287 LQOKSYKGDPIVQHRPMLBEIYELFRLMKSAFVEROPCMHPRPPIVYIKTVQFT 346  
 QY 241 TKVRLVKEPELNYQKIKVCIDKSGDVAA 271  
 DB 347 TKVRLVKEPELNYQKIKVCIDKSGDVAA 377  
 RESULT 4  
 AAE22055 standard; protein; 720 AA.  
 ID AAE22055;  
 AC AAE22055;  
 XX  
 XX 25-JUL-2002 (first entry)  
 DT  
 XX  
 XX Human Stat3beta protein.  
 DE  
 XX  
 XX Human; signal transducer and activator of transcription 3; ischaemia;  
 KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;  
 KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;  
 KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;  
 KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;  
 KW shock; chronic active hepatitis; adult respiratory distress syndrome;  
 KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;  
 KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;  
 KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;  
 KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;  
 KW carcinoma; degenerative disorder; gene therapy; growth deficiency;  
 KW cirrhosis; hypoproliferative disorder; lesion; statbeta.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers

FT Misc-difference 713. 714  
 FT /note="Encoded by ACA CCA TTC"  
 XX  
 XX WO200220032-A1.  
 XX  
 XX  
 PD 14-MAR-2002.  
 XX  
 XX 10-SEP-2001; 2001WO-US028254.  
 XX  
 XX 08-SEP-2000; 2000US-0231212P.  
 XX  
 XX (UYJO ) UNIV JOHNS HOPKINS.  
 XX (UYF-) UNIV SOUTH FLORIDA.  
 XX  
 XX Yu H, Pardoll D, Jove R, Dalton W;  
 XX WPI, 2002-362218/39.  
 DR N-PSDB; AAD35066.  
 XX  
 XX  
 PT Modulating angiogenesis and an immune response in an individual, for  
 PT treating a hypoxic or ischemic condition, comprises administering a  
 PT compound that modulates the activity of a signal transducer and activator  
 PT of transcription 3.  
 XX  
 XX  
 PS Disclosure; Page 87-89; 94pp; English.  
 XX  
 XX  
 CC The invention relates to a method of modulating angiogenesis and immune  
 CC response. Method involves administering to an individual a compound that  
 CC modulate the activity of signal transducer and activator of transcription  
 CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing  
 CC hypoxic or ischemic condition or disorder which is the result of stroke,  
 CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,  
 CC tissue ischaemia in the lower extremities, infarction, trauma, vascular  
 CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,  
 CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,  
 CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,  
 CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy  
 CC with neovascularisation. Suppressing an immune response is useful for  
 CC ameliorating a symptom of an autoimmune disease such as systemic lupus  
 CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,  
 CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,  
 CC mixed connective tissue disease, primary biliary cirrhosis, pernicious  
 CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,  
 CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,  
 CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's  
 CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune  
 CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and  
 CC dense deposit disease. The method is useful in preventing or treating  
 CC specific proliferative and oncogenic disease which includes sarcomas and  
 CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,  
 CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,  
 CC hypoproliferative disorders, physical trauma, lesions and wounds. The  
 CC method is also used in gene therapy. The present sequence is human  
 CC Stat3beta protein  
 CC  
 SQ Sequence 720 AA;  
 Query Match 99.3%; Score 1377; DB 5; Length 720;  
 Best Local Similarity 99.3%; Pred. No. 9.1e-115;  
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RCLMESRLDTAATTAAGGGAHPTAAVTEKQOMLEOHLDVREKVDLEQKAYVE 60  
 DB 107 RCLMESRLDTAATTAAGGGAHPTAAVTEKQOMLEOHLDVREKVDLEQKAYVE 166  
 QY 61 NLQDDFDNRYKTKSGQGMODLNGNSVTRQKMOLEQMTALDQMRSTVSELAGLLS 120  
 DB 167 NLQDDFDNRYKTKSGQGMODLNGNSVTRQKMOLEQMTALDQMRSTVSELAGLLS 226  
 QY 121 AMEVYQKTLTBEALADMKRRPPIACIGPPNICDLRLNMTTSLAESQLQTRQOIKKLE 180  
 DB 227 AMEVYQKTLTBEALADMKRRPPIACIGPPNICDLRLNMTTSLAESQLQTRQOIKKLE 286

QY 181 LQOKVSYKDDPIVOHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 240  
 DB 287 LQOKVSYKDDPIVOHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 346  
 QY 241 TKVRLVKFPELNYOLKIVCICDKSGDVAA 271  
 DB 347 TKVRLVKFPELNYOLKIVCICDKSGDVAA 377

## RESULT 5

ABB57164  
 ID ABB57164 standard; protein; 769 AA.

AC ABB57164;  
 XX

DT 07-MAR-2002 (first entry)  
 XX

DE Mouse ischemic condition related protein sequence SEQ ID NO:398.  
 XX

KW Mouse; ischemia; compressive ischemia; occlusive ischemia;  
 KM vasospastic ischemia; ischemic condition; ischemic disease.  
 XX

OS Mus musculus.  
 XX

PN WO200188188-A2.  
 XX

PD 22-NOV-2001.  
 XX

PF 18-MAY-2001; 2001WO-JP04192.  
 XX

PR 18-MAY-2000; 2000JP-00145977.  
 XX

PA (UNIV-) UNIV NITON SCHOOL JURIDICAL PERSON.  
 XX

PI Ishikawa K, Arai S, Takahashi Y, Nagata T, Ishii Y;  
 XX

DR WPI, 2002-034733/04.  
 XX

DR N-PSDB; ABI99454.  
 XX

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
 PT expression levels of particular genes defined in the specification or by  
 PT determining the expression profile of a gene group comprising these  
 PT genes.

Claim 2; Page 1084-1087; 2690pp; English.

CC The present invention describes a method for examining ischemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC gene (1) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (1). The method  
 CC is useful for examining the ischemic condition (e.g. compressive  
 CC ischemia, occlusive ischemia or vasospastic ischemia) by measuring the  
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischemic condition-improving drugs or  
 CC therapeutics for ischemic diseases. ABI99913 and ABI99914 represent PCR  
 CC primers for a mouse ischemic condition related sequence, which are used  
 CC in the exemplification of the present invention  
 CC

XX Sequence 769 AA;  
 SQ

Query Match 99.2%; Score 1377; DB 5; Length 769;  
 Best Local Similarity 99.3%; Pred. No. 1e-114;

Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMESRLQTRATTAAGGAGNHTAAVTEKQMLROHLDVPRKRVQDLEQKKVYE 60  
 DB 107 RCLMESRLQTRATTAAGGAGNHTAAVTEKQMLROHLDVPRKRVQDLEQKKVYE 166  
 QY 61 NLODDPDEFYKTLKSGQDMQDINGNNSVTRQKMOQLTALDQMRBSIVSELAGLIS 120

DB 167 NLODDPDEFYKTLKSGQDMQDINGNNSVTRQKMOQLTALDQMRBSIVSELAGLIS 226  
 QY 121 AMEYVOKTITDEBLAMKRRPEIACGPPNICLDRLENWITSLAESQLOTRQOIKLEE 180  
 DB 227 AMEYVOKTITDEBLAMKRRQOACIGPPNICLDRLENWITSLAESQLOTRQOIKLEE 286  
 QY 181 LQOKVSYKDDPIVOHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 240  
 DB 287 LQOKVSYKDDPIVOHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 346  
 QY 241 TKVRLVKFPELNYOLKIVCICDKSGDVAA 271  
 DB 347 TKVRLVKFPELNYOLKIVCICDKSGDVAA 377

## RESULT 6

AAE22054  
 ID AAE22054 standard; protein; 769 AA.

AC AAE22054;  
 XX

DT 25-JUL-2002 (first entry)  
 XX

DE Human Stat3 protein.  
 XX

KW Human; signal transducer and activator of transcription 3; ischemia;  
 KM immune response; Stat3; coronary atherosclerosis; vascular occlusion;  
 KM hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;  
 KM inflammation; chronic obstructive pulmonary disease; cardiac arrest;  
 KM insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;  
 KM shock; chronic active hepatitis; adult respiratory distress syndrome;  
 KM nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;  
 KM Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;  
 KM polyomyelitis; rheumatoid arthritis; autoimmune infertility; anaemia;  
 KM proliferative disease; Grave's disease; ulcerative colitis; sarcoma;  
 KM carcinoma; degenerative disorder; gene therapy; growth deficiency;  
 KM cirrhosis; hypoproliferative disorder; lesion.  
 XX

OS Homo sapiens.  
 XX

PN WO200220032-A1.  
 XX

PD 14-MAR-2002.  
 XX

PF 10-SEP-2001; 2001WO-US028254.  
 XX

PR 08-SEP-2000; 2000US-0231212P.  
 XX

PA (UNIV-) UNIV JOHNS HOPKINS.  
 XX

PI (USF-) UNIV SOUTH FLORIDA.  
 XX

PI Yu H, Pardoll D, Jove R, Dalton W;  
 XX

DR WPI; 2002-362218/39.  
 XX

DR N-PSDB; AAD35065.  
 XX

PT Modulating angiogenesis and an immune response in an individual, for  
 PT treating a hypoxic or ischemic condition, comprises administering a  
 PT compound that modulates the activity of a signal transducer and activator  
 PT of transcription 3.  
 PT

PS Disclosure; Page 83-85; 94pp; English.

CC The invention relates to a method of modulating angiogenesis and immune  
 CC response. Method involves administering to an individual a compound that  
 CC modulate the activity of signal transducer and activator of transcription  
 CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing  
 CC hypoxic or ischemic condition or disorder which is the result of stroke,  
 CC ischemia, coronary atherosclerosis, myocardial infarction, inflammation,  
 CC tissue ischemia in the lower extremities, infarction, trauma, vascular  
 CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,  
 CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,  
 CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest.

nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy  
 CC with neovascularisation. Suppressing an immune response is useful for  
 ameliorating a symptom of an autoimmune disease such as systemic lupus  
 erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,  
 Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,  
 mixed connective tissue disease, primary biliary cirrhosis, pernicious  
 anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,  
 gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,  
 idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's  
 disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune  
 infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and  
 dense deposit disease. The method is useful in preventing or treating  
 specific proliferative and oncogenic disease which includes sarcomas and  
 carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,  
 fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,  
 CC hypoproliferative disorders, physical trauma, lesions and wounds. The  
 CC method is also used in gene therapy. The present sequence is human Stat3  
 CC protein

XX  
 XX  
 SQ Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;

Beet Local Similarity 99.3%; Pred. No. 1e-114; Mismatches 1; Indels 0; Gaps 0;

Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMEESRLQTPAATAAQQGQANPTAAVTEKQOMLEOHLDVRRKRVODLEQKMKVE 60  
 DB 107 RCLMEESRLQTPAATAAQQGQANPTAAVTEKQOMLEOHLDVRRKRVODLEQKMKVE 166

QY 61 NIQDDPFRNYKTKSQGQMODLNGNNSVTRQKQOLEOMLTALDQMRRSYSELAGLIS 120  
 DB 167 NIQDDPFRNYKTKSQGQMODLNGNNSVTRQKQOLEOMLTALDQMRRSYSELAGLIS 226

QY 121 AMEYVQKLTLDLBEADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOIKLEE 180  
 DB 227 AMEYVQKLTLDLBEADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOIKLEE 286

QY 181 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCMRPHRPRPLVTKGVQPT 240  
 DB 287 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCMRPHRPRPLVTKGVQPT 346

QY 241 TKVRLVFPPELNYQKTKVCIDKSGDVAA 271  
 DB 347 TKVRLVFPPELNYQKTKVCIDKSGDVAA 377

RESULT 7

AAE22056  
 ID AAE22056 standard; protein; 769 AA.

XX  
 AC AAE22056;

XX  
 DT 25-JUL-2002 (first entry)

XX  
 DE Human protein related to angiogenesis regulation.

XX  
 KM Human; signal transducer and activator of transcription 3; ischaemia;  
 KM immune response; Stat3; coronary atherosclerosis; vascular occlusion;  
 KM hyposmia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;  
 KM inflammation; chronic obstructive pulmonary disease; cardiac arrest;  
 KM insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;  
 KM shock; chronic active hepatitis; adult respiratory distress syndrome;  
 KM nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;  
 KM Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;  
 KM polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;  
 KM proliferative disease; Grave's disease; ulcerative colitis; sarcoma;  
 KM carcinoma; degenerative disorder; gene therapy; growth deficiency;  
 KM cirrhosis; hypoproliferative disorder; lesion.

XX  
 OS Homo sapiens.

XX  
 PN WO200220032-A1.

PD 14-MAR-2002.  
 XX  
 XX 10-SEP-2001; 2001WO-US028254.  
 PF  
 XX  
 PR 08-SEP-2000; 2000US-0231212P.  
 XX  
 XX (UYVO ) UNIV JOHNS HOPKINS.  
 PA (UYSF-) UNIV SOUTH FLORIDA.  
 XX  
 XX Yu H, Pardoll D, Jove R, Dalton W;  
 PI WPI; 2002-362218/39.  
 DR  
 XX  
 XX  
 PT Modulating angiogenesis and an immune response in an individual, for  
 PT treating a hypoxic or ischemic condition, comprises administering a  
 PT compound that modulates the activity of a signal transducer and activator  
 PT of transcription 3.

PS Disclosure; Page 83-85; 94pp; English.

XX  
 XX  
 CC The invention relates to a method of modulating angiogenesis and immune  
 CC response. Method involves administering to an individual a compound that  
 CC modulate the activity of signal transducer and activator of transcription  
 CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing  
 CC hypoxic or ischemic condition or disorder which is the result of stroke,  
 CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,  
 CC tissue ischaemia in the lower extremities, infarction, trauma, vascular  
 CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,  
 CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,  
 CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,  
 CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy  
 CC with neovascularisation. Suppressing an immune response is useful for  
 CC ameliorating a symptom of an autoimmune disease such as systemic lupus  
 CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,  
 CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,  
 CC mixed connective tissue disease, primary biliary cirrhosis, pernicious  
 CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,  
 CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,  
 CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's  
 CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune  
 CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and  
 CC dense deposit disease. The method is useful in preventing or treating  
 CC specific proliferative and oncogenic disease which includes sarcomas and  
 CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,  
 CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,  
 CC hypoproliferative disorders, physical trauma, lesions and wounds. The  
 CC method is also used in gene therapy. The present sequence is human  
 CC protein related to angiogenesis regulation

XX  
 SQ Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;

Beet Local Similarity 99.3%; Pred. No. 1e-114; Mismatches 1; Indels 0; Gaps 0;

Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMEESRLQTPAATAAQQGQANPTAAVTEKQOMLEOHLDVRRKRVODLEQKMKVE 60  
 DB 107 RCLMEESRLQTPAATAAQQGQANPTAAVTEKQOMLEOHLDVRRKRVODLEQKMKVE 166

QY 61 NIQDDPFRNYKTKSQGQMODLNGNNSVTRQKQOLEOMLTALDQMRRSYSELAGLIS 120  
 DB 167 NIQDDPFRNYKTKSQGQMODLNGNNSVTRQKQOLEOMLTALDQMRRSYSELAGLIS 226

QY 121 AMEYVQKLTLDLBEADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOIKLEE 180  
 DB 227 AMEYVQKLTLDLBEADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOIKLEE 286

QY 181 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCMRPHRPRPLVTKGVQPT 240  
 DB 287 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCMRPHRPRPLVTKGVQPT 346

QY 241 TKVRLVFPPELNYQKTKVCIDKSGDVAA 271  
 |||

Db 347 TKVRLVKEPELNYOLKIKVCIDKDSGDVAA 377

RESULT 8  
AAR82995

ID AAR82995 standard; protein; 770 AA.

AC AAR82995;

DT 25-MAR-1996 (first entry)

DE Mouse liver acute phase response factor.

XX Mouse; acute phase response factor; transcription factor; interleukin-6;  
XX signal transduction; liver; antibody; antisense; ribozyme;  
XX antiinflammatory; antitumor; hypotensive; therapy.

XX Mus musculus.

OS Mus musculus.

PN EP676469-A2.

PD 11-OCT-1995.

PF 23-MAR-1995; 95EP-00104670.

PR 04-APR-1994; 94UP-00065825.

PA (KISH/) KISHIMOTO T.

PI Akira S, Kishimoto T;

XX MPI; 1995-346089/45.

DR N-PSDB; AAT05619.

PT New acute phase response factor - for developing inhibitory agents for  
PT treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory  
PT diseases.

XX Claim 10; Page 20-22; 31pp; English.

XX The sequence represents a mouse acute phase response factor (APRF), a  
XX transcription factor related to signal transduction of interleukin-6 (IL-  
XX 6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA  
XX library using a polymerase chain reaction product (amplified using  
XX primers derived from an IL-6-treated mouse liver peptide) as a probe.  
XX APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or  
XX ribozymes, may be used to treat diseases induced by IL-6, e.g.  
XX inflammatory disease, leukemia, cancer, osteoclastia, pulmonary  
XX hypertension, etc

SQ Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 2; Length 770;  
Best Local Similarity 99.3%; Pred. No. 1e-114; Indels 0; Gaps 0;  
Matches 269; Conservative 1; Mismatches 1;

QY 1 RCIMESRLIQTATATAAAGGQANHPPTAAVTEKQOMLEOHLODVKKRVODLEQKKVVE 60  
DB 107 RCIMESRLIQTATATAAAGGQANHPPTAAVTEKQOMLEOHLODVKKRVODLEQKKVVE 166

QY 61 NIODDDPFNYKTLKSGDMODLNGNNSVTRKMOQLEQMLTALDQMRISIVSELGLLS 120  
DB 167 NIODDDPFNYKTLKSGDMODLNGNNSVTRKMOQLEQMLTALDQMRISIVSELGLLS 226

QY 121 AMEYVOKTLDEBLADMKRPRPIACIGPPNICLDRLENNITSLASQLOTRQOIKLEE 180  
DB 227 AMEYVOKTLDEBLADMKRPRPIACIGPPNICLDRLENNITSLASQLOTRQOIKLEE 286

QY 181 LQOKSVKGDPIVOHRPMLERIVELFRNIMKSAFVEROPCMRPHDRDLVITKGVQFT 240  
DB 287 LQOKSVKGDPIVOHRPMLERIVELFRNIMKSAFVEROPCMRPHDRDLVITKGVQFT 346

QY 241 TKVRLVKEPELNYOLKIKVCIDKDSGDVAA 271

Db 347 TKVRLVKEPELNYOLKIKVCIDKDSGDVAA 377

RESULT 9  
AAY03768

ID AAY03768 standard; protein; 770 AA.

AC AAY03768;

DT 11-JUN-1999 (first entry)

DE Human STAT3 allelic variant.

XX Signal Transducer and Activator of Transcription 3 (STAT3) allele; IL-6;  
XX intracellular transcription factor; interleukin-6; medicament; variant;  
XX pharmaceutical; autoimmune disease; inflammatory; human.

XX Homo sapiens.

PN EP905234-A2.

PD 31-MAR-1999.

PF 18-FEB-1998; 98EP-00102774.

PR 16-SEP-1997; 97EP-00116061.

PA (ISTF) ARI APPLIED RES SYSTEMS HOLDING NV.

PI Serlupi-Crescenzi O, Della Pietra L;

XX MPI; 1999-192664/17.

DR N-PSDB; AAX29281.

PT New human Signal Transducer and Activator of Transcription 3 (STAT3)  
PT allelic variant useful for treatment of autoimmune and inflammatory  
PT disease.

XX Claim 2; Page 9-13; 32pp; English.

XX The present sequence represents a predominant allelic variant of human  
XX Signal Transducer and Activator of Transcription 3 (STAT3) protein, an  
XX intracellular transcription factor which mediates IL-6 signals. The  
XX encoding sequence differs from the original published human STAT3 gene  
XX sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3  
XX DNA molecule can be used for the recombinant expression of the variant.  
XX STAT3 protein is useful as a medicament or pharmaceutical composition for  
XX treatment of autoimmune or inflammatory diseases

SQ Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 2; Length 770;  
Best Local Similarity 99.3%; Pred. No. 1e-114; Indels 0; Gaps 0;  
Matches 269; Conservative 1; Mismatches 1;

QY 1 RCIMESRLIQTATATAAAGGQANHPPTAAVTEKQOMLEOHLODVKKRVODLEQKKVVE 60  
DB 107 RCIMESRLIQTATATAAAGGQANHPPTAAVTEKQOMLEOHLODVKKRVODLEQKKVVE 166

QY 61 NIODDDPFNYKTLKSGDMODLNGNNSVTRKMOQLEQMLTALDQMRISIVSELGLLS 120  
DB 167 NIODDDPFNYKTLKSGDMODLNGNNSVTRKMOQLEQMLTALDQMRISIVSELGLLS 226

QY 121 AMEYVOKTLDEBLADMKRPRPIACIGPPNICLDRLENNITSLASQLOTRQOIKLEE 180  
DB 227 AMEYVOKTLDEBLADMKRPRPIACIGPPNICLDRLENNITSLASQLOTRQOIKLEE 286

QY 181 LQOKSVKGDPIVOHRPMLERIVELFRNIMKSAFVEROPCMRPHDRDLVITKGVQFT 240  
DB 287 LQOKSVKGDPIVOHRPMLERIVELFRNIMKSAFVEROPCMRPHDRDLVITKGVQFT 346

QY 241 TKVRLVKEPELNYOLKIKVCIDKDSGDVAA 271

Db 347 TKVRLVKEPPELNYOLKIKVCIDKSGDVAA 377

RESULT 10  
AAB12377

ID AAB12377 standard; peptide; 770 AA.

AC AAB12377;

DT 08-NOV-2000 (first entry)

DE N-terminal domain of murine STAT-3 protein.

KW STAT: signal transducer and activator of transcription; crystal;  
drug design; murine.

OS Mus sp.

Key Location/Qualifiers

FT 4..9 /label= Alpha helix 1

FT 12..21 /label= Alpha helix 2

FT 19..21 /label= 3(10) helix of alpha helix 2

FT 28..33 /label= Alpha helix 3

FT 35..40 /label= Alpha helix 4

FT 43..47 /label= Alpha helix 5

FT 50..73 /label= Alpha helix 6

FT 77..96 /label= Alpha helix 7

FT 99..119 /label= Alpha helix 8

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

Best Local Similarity 99.3%; Pred. No. 1e-114; Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCIMESRLQTAATPAAGGQANHPATAVTEKQMLEOHLODYRKXVQDLEOKMYVE 60

Db 107 RCLMESRLQTAATPAAGGQANHPATAVTEKQMLEOHLODYRKXVQDLEOKMYVE 166

QY 61 NLQDDFDFNYKTKLSQGMQDNLGNNQSVTRQKMQLEQMTALDQMRSSIVSELAGILS 120

Db 167 NLQDDFDFNYKTKLSQGMQDNLGNNQSVTRQKMQLEQMTALDQMRSSIVSELAGILS 226

QY 121 AMEYVQKTLTDEELADMRKRPPEIACIGPPNICDRLNMTISLASQLOTRQOIKLEE 180

Db 227 AMEYVQKTLTDEELADMRKRPPEIACIGPPNICDRLNMTISLASQLOTRQOIKLEE 286

QY 181 LQOKVSYKGDPIVQHPMLERIVELEFNIMKSAFVVEROPCPMPHPDPPLVYKIGVPT 240

Db 287 LQOKVSYKGDPIVQHPMLERIVELEFNIMKSAFVVEROPCPMPHPDPPLVYKIGVPT 346

QY 241 TKVRLVKEPPELNYOLKIKVCIDKSGDVAA 271

Db 347 TKVRLVKEPPELNYOLKIKVCIDKSGDVAA 377

RESULT 11

ID AAE14652 standard; protein; 770 AA.

AC AAE14652;

DT 16-JUL-2002 (first entry)

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

Best Local Similarity 99.3%; Pred. No. 1e-114; Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCIMESRLQTAATPAAGGQANHPATAVTEKQMLEOHLODYRKXVQDLEOKMYVE 60

Db 107 RCLMESRLQTAATPAAGGQANHPATAVTEKQMLEOHLODYRKXVQDLEOKMYVE 166

QY 61 NLQDDFDFNYKTKLSQGMQDNLGNNQSVTRQKMQLEQMTALDQMRSSIVSELAGILS 120

Db 167 NLQDDFDFNYKTKLSQGMQDNLGNNQSVTRQKMQLEQMTALDQMRSSIVSELAGILS 226

QY 121 AMEYVQKTLTDEELADMRKRPPEIACIGPPNICDRLNMTISLASQLOTRQOIKLEE 180

Db 227 AMEYVQKTLTDEELADMRKRPPEIACIGPPNICDRLNMTISLASQLOTRQOIKLEE 286

QY 181 LQOKVSYKGDPIVQHPMLERIVELEFNIMKSAFVVEROPCPMPHPDPPLVYKIGVPT 240

Db 287 LQOKVSYKGDPIVQHPMLERIVELEFNIMKSAFVVEROPCPMPHPDPPLVYKIGVPT 346

QY 241 TKVRLVKEPPELNYOLKIKVCIDKSGDVAA 271

Db 347 TKVRLVKEPPELNYOLKIKVCIDKSGDVAA 377

RESULT 11

ID AAE14652 standard; protein; 770 AA.

AC AAE14652;

DT 16-JUL-2002 (first entry)

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.

DE Murine STAT3 protein.



CC the control of a promoter containing at least two adjacent weak binding  
 CC sites for STAT protein dimers. The methods are used for identifying new  
 CC drugs. An antagonist of STAT N-terminal dimeric interactions that  
 CC inhibits the binding of the STAT dimers to adjacent weak binding sites on  
 CC a promoter of a gene, could be useful as drugs in the treatment of  
 CC diseases, e.g. inflammation, allergy, asthma and leukaemia. On the other  
 CC hand, an agonist of N-terminal dimeric interactions between STAT dimers,  
 CC can be used as drugs in the treatment of diseases e.g. anaemia,  
 CC neutropenia, thrombocytopaenia, cancer, obesity, viral diseases and  
 CC growth retardation. The present sequence is murine STAT3 protein

XX Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 5; Length 770;  
 Best Local Similarity 99.3%; Pred. No. 1e-114;  
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMESRLIQTATAAOGGQANHPAAVTEKQMLDEHLDVRRVODLEQKKVVE 60  
 Db 107 RCLMESRLIQTATAAOGGQANHPAAVTEKQMLDEHLDVRRVODLEQKKVVE 166  
 QY 61 NLDDDFPNYKTKSGQDMQDNGNNSVTRQKQMLEQMLTALDQRRSIVSELGILLS 120  
 Db 167 NLDDDFPNYKTKSGQDMQDNGNNSVTRQKQMLEQMLTALDQRRSIVSELGILLS 226  
 QY 121 AMEYVQKTLTDEELADMKRRPELACTGGPNNICLDLENNWITSLASSQLQTRQOIKLEE 180  
 Db 227 AMEYVQKTLTDEELADMKRRPELACTGGPNNICLDLENNWITSLASSQLQTRQOIKLEE 286  
 QY 181 LOOKVSYKGDPIYQHRPMLERIVELFRNLMSAFVVERPCMPMPDRPLVTKTGQFT 240  
 Db 287 LOOKVSYKGDPIYQHRPMLERIVELFRNLMSAFVVERPCMPMPDRPLVTKTGQFT 346  
 QY 241 TKVRLIVKPEELNYQLKIKVCIDKSGDVAA 271  
 Db 347 TKVRLIVKPEELNYQLKIKVCIDKSGDVAA 377

RESULT 12

ABG69497 ID ABG69497 standard; protein; 770 AA.

XX ABG69497;

XX 21-OCT-2002 (first entry)

XX Human bait protein STAT3.

XX Human; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;

XX non-insulin diabetes mellitus; obesity; selected interacting domain; SID,  
 XX protein-protein interaction map; PIM; anorectic; metabolic disorder.

XX Homo sapiens.

XX W0200253726-A2.

XX 11-UTL-2002.

XX 28-DEC-2001; 2001WO-EP015423.

XX 02-JAN-2001; 2001US-02599377P.

XX (HYBR-) HYBRIGENICS.

XX (CNRS) CENT NAT RECH SCI.

XX Legrain P, Marullo S, Jockers R;

XX WPI; 2002-583612/62.

XX N-PSDB; ABS51033.

XX Novel complex of protein-protein interactions in adipocyte cells for  
 PT identifying compounds that modulate the protein-protein interactions and  
 PT useful for treating obesity and metabolic disorders.

XX Claim 1; Page 54; 125pp; English.

XX The invention relates to a complex of protein-protein interactions  
 CC (forming a protein-protein interaction map, PIM) in adipocyte cells as  
 CC defined in the specification, or polynucleotides in adipocytes encoding the  
 CC for the polypeptides. Also included are a recombinant cell expressing the  
 CC interacting polypeptides and a method of selecting a modulating compound  
 CC in adipocyte cells, by culturing a recombinant host cell on a selective  
 CC medium containing a modulating compound and a reporter gene the  
 CC expression of which is toxic for the recombinant host cell which is  
 CC transformed with two vectors, where the first vector comprises a  
 CC polynucleotide encoding a first hybrid polypeptide and DNA binding domain  
 CC and the second vector comprising a polynucleotide encoding a second  
 CC hybrid polypeptide and an activating domain that activates the toxic  
 CC reporter gene, when the first and second hybrid polypeptides interact and  
 CC selecting the modulating compound which inhibits the growth of the  
 CC recombinant host cell (i.e. using the yeast two-hybrid system). The  
 CC complexes are useful for identifying compounds that modulate the protein-  
 CC protein interactions and useful for treating obesity and metabolic  
 CC disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The  
 CC compound isolated by the method is useful for treating and preventing  
 CC obesity or metabolic diseases. The interactions between the proteins of  
 CC the complex further define a set of selected interacting domains, SID.  
 CC The present sequence represents a member of the protein complex of the  
 CC invention, used as the bait protein in the yeast two-hybrid assay

XX Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 5; Length 770;  
 Best Local Similarity 99.3%; Pred. No. 1e-114;  
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMESRLIQTATAAOGGQANHPAAVTEKQMLDEHLDVRRVODLEQKKVVE 60  
 Db 107 RCLMESRLIQTATAAOGGQANHPAAVTEKQMLDEHLDVRRVODLEQKKVVE 166  
 QY 61 NLDDDFPNYKTKSGQDMQDNGNNSVTRQKQMLEQMLTALDQRRSIVSELGILLS 120  
 Db 167 NLDDDFPNYKTKSGQDMQDNGNNSVTRQKQMLEQMLTALDQRRSIVSELGILLS 226  
 QY 121 AMEYVQKTLTDEELADMKRRPELACTGGPNNICLDLENNWITSLASSQLQTRQOIKLEE 180  
 Db 227 AMEYVQKTLTDEELADMKRRPELACTGGPNNICLDLENNWITSLASSQLQTRQOIKLEE 286  
 QY 181 LOOKVSYKGDPIYQHRPMLERIVELFRNLMSAFVVERPCMPMPDRPLVTKTGQFT 240  
 Db 287 LOOKVSYKGDPIYQHRPMLERIVELFRNLMSAFVVERPCMPMPDRPLVTKTGQFT 346  
 QY 241 TKVRLIVKPEELNYQLKIKVCIDKSGDVAA 271  
 Db 347 TKVRLIVKPEELNYQLKIKVCIDKSGDVAA 377

RESULT 13

ABU10476 ID ABU10476 standard; protein; 770 AA.

XX ABU10476;

XX 06-AUG-2003 (first entry)

XX Mouse STAT3 protein.

XX Mouse; signal transducer and activator of transcription; drug design;  
 XX screening; STAT-STAT dimer interaction; STAT3.

XX Mus sp.

XX Key

XX Region

XX Location/Qualifiers  
 FT 4..9  
 FT /label= alpha\_helix\_1  
 FT 12..21  
 FT Region



```

FT FT /label= alpha_helix_2
FT FT /note= "Residues 19-21 form a 3 helix"
FT FT 28..33
FT FT /label= alpha_helix_3
FT FT 35..40
FT FT /label= alpha_helix_4
FT FT 43..47
FT FT /label= alpha_helix_5
FT FT 50..73
FT FT /label= alpha_helix_6
FT FT /note= "Residues 57, 61, 64, 68 and 71 contribute to
FT FT packing of the coiled-coil"
FT FT 77..96
FT FT /label= alpha_helix_7
FT FT /note= "Residues 79, 83, 86, 90 and 94 contribute to
FT FT packing of the coiled-coil"
FT FT 99..119
FT FT /label= alpha_helix_8
FT FT
FT FT Region
FT FT
FT FT US2003003563-A1.
FT FT
FT FT 02-JAN-2003.
FT FT
FT FT 19-OCT-2001; 2001US-00045792.
FT FT
FT FT 23-JAN-1998; 98US-00012710.
FT FT 24-APR-2000; 2000US-00556273.
FT FT
FT FT (VINK/) VINKMEIER U.
FT FT (MOAR/) MOAREFI I.
FT FT (DARN/) DARNELL J E.
FT FT (KURI/) KURIYAN J.
FT FT
FT FT Vinmeier U, Moarefi I, Darnell JE, Kuriyan J;
FT FT
FT FT WPI; 2003-447354/42.
FT FT
FT FT New crystal having an N-terminal domain of a STAT protein performing X-
FT FT ray crystallographic studies, useful for screening drugs that enhance or
FT FT inhibit STAT-STAT dimer interactions.
FT FT
FT FT Disclosure; Page 25-26; 46pp; English.
FT FT
FT FT The invention relates to a crystal of an N-terminal domain of signal
FT FT transducer and activator of transcription (STAT) protein, where the
FT FT crystal effectively diffracts X-rays for the determination of the atomic
FT FT coordinates of the N-terminal domain of the STAT protein to a resolution
FT FT of greater than 5.0 Angstrom. The methods and compositions are useful for
FT FT the design and screening of drugs that enhance or inhibit STAT-STAT dimer
FT FT interactions. The present sequence represents the amino acid sequence of
FT FT mouse STAT3 protein
FT FT
FT FT Sequence 770 AA;
FT FT
FT FT Query Match 99.2%; Score 1377; DB 6; Length 770;
FT FT Best Local Similarity 99.3%; Pred. No. 1e-114;
FT FT Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 241 TKVRLVKEPELNVQLKIKVCIDKSGDVAA 271
DB 347 TKVRLVKEPELNVQLKIKVCIDKSGDVAA 377
DB
RESULT 14
ADN04365
ID ADN04365 standard; protein; 770 AA.
XX
XX ADN04365;
XX
XX 01-JUL-2004 (first entry)
XX
XX Antipsoriatic protein sequence #377.
DE
XX Antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
XX antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
XX Homo sapiens.
OS
XX
XX WO2004028479-A2.
XX
XX 08-APR-2004.
XX
XX 25-SEP-2003; 2003WO-US030907.
XX
XX 25-SEP-2002; 2002US-0414006P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX PI Wu JD;
XX
XX WPI; 2004-305105/28.
XX
XX N-PSDB; ADN04364.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX
XX Claim 9; SEQ ID NO 759; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polypeptides of the invention.
XX
XX Sequence 770 AA;
XX
XX Query Match 99.2%; Score 1377; DB 8; Length 770;
XX Best Local Similarity 99.3%; Pred. No. 1e-114;
XX Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

RESULT 15  
AAB58442  
ID AAB58442 standard; protein; 793 AA.  
XX  
AC AAB58442;  
DT 14-MAR-2001 (first entry)  
XX  
DE Lung cancer associated polypeptide sequence SEQ ID 780.  
XX  
KW Human; lung cancer associated protein; neuroprotective; cyrostatic;  
KW cardioactive; immunomodulatory; muscular active; vulnerary;  
KW gastrointestinal; nephrotoxic; antiinfective; gynecological;  
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
KW proliferative disorder; wound healing; infectious disease.  
XX  
OS Homo sapiens.  
XX  
PN WO20005180-A2.  
XX  
PD 21-SEP-2000.  
XX  
PF 08-MAR-2000; 2000MO-US005918.  
XX  
PR 12-MAR-1999; 99US-0124270P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI (ROSE/) ROSEN C A.  
XX  
PI Ruben SM;  
XX  
DR WPI; 2000-587514/55.  
DR N-PSDB; AAF18318.  
XX  
PT Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer.  
XX  
PS Claim 11; Page 1310-1313; 1425pp; English.  
XX  
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective; cyrostatic; cardioactive;  
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal  
CC general; nephrotoxic; antiinfective; gynecological; or antibacterial  
CC activity. The invention also includes antibodies specific for the protein  
CC or polynucleotide sequences. The lung cancer associated polynucleotide  
CC sequences may be used for detection of lung cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The proteins may be used to treat disorders such as  
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,  
CC cardiovascular, renal, and proliferative disorders. The proteins may also  
CC be used in the treatment of wounds and infectious diseases.  
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are  
CC used in the course of the invention for the identification and  
CC characterisation of the polynucleotide and protein sequences  
XX  
SQ Sequence 793 AA:

Query Match 99.2%; Score 1377; DB 3; Length 793;  
Best Local Similarity 99.3%; Pred. No. 1e-114;  
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMBESRLLOQTATTAAGGAGANHPAAVTEKQMLBOHLDVRRKRVODLEOKMKVVE 60  
DB 130 RCLMBESRLLOQTATTAAGGAGANHPAAVTEKQMLBOHLDVRRKRVODLEOKMKVVE 189  
QY 61 NLQDDDFPNYKTLKSGDMQDLNGNNSVTRQKQOLEQMLTALDQMRSSIVSELGLLS 120  
DB 190 NLQDDDFPNYKTLKSGDMQDLNGNNSVTRQKQOLEQMLTALDQMRSSIVSELGLLS 249

QY 121 AMEYVQKTLTDEELADWKRREPEIACIGPPNIIQDRLENNITSLAESQLOTRQIKKLE 180  
DB 250 AMEYVQKTLTDEELADWKRREPEIACIGPPNIIQDRLENNITSLAESQLOTRQIKKLE 309  
QY 181 LQKVSYSKDPPIVOHRPMLBERIVELFRNLKSAFVERQPCMPHDPDRPIVITGVQFT 240  
DB 310 LQKVSYSKDPPIVOHRPMLBERIVELFRNLKSAFVERQPCMPHDPDRPIVITGVQFT 369  
QY 241 TKVRLVKKPELNYOLKIKVCIDKSGDVAA 271  
DB 370 TKVRLVKKPELNYOLKIKVCIDKSGDVAA 400

Search completed: December 2, 2004, 23:49:06  
Job time : 161 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: December 2, 2004, 23:35:28 ; Search time 193 Seconds

(without alignments)  
807.909 Million cell updates/sec

Title: US-10-090-185-9  
Perfect score: 1388  
Sequence: 1 RCLMEBSRLQTAATAAQQG.....LNYQIKIKVICIDKSGDVAA 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1377	99.2	722	AAH19168	Aah19168 mus muscu
2	1377	99.2	770	STAX3_HUMAN	P40763 homo sapien
3	1377	99.2	770	STAX3_MOUSE	P44227 mus musculi
4	1374	99.0	770	STAX3_RAT	P56231 rattus norv
5	1370	98.7	770	STAX3_BOVIN	P61635 bos taurus
6	1352	97.4	771	O6DV79	O6dv79 gallus gall
7	1304	93.9	769	O9PVX8	O9pvx8 xenopus lae
8	1286	92.7	766	O7ZKX3	O7ztx3 xenopus lae
9	1191.5	85.8	414	O7ZTS5	O7zts5 brachydanto
10	1191.5	85.8	786	O6WV46	O6wv46 brachydanto
11	1191.5	85.8	786	AAH68320	Aah68320 brachydanto
12	1191.5	85.8	806	O93599	O93599 brachydanto
13	1182.5	85.2	765	O6QVE3	O6qv3 oryzias lat
14	1182.5	85.2	785	O6QVE7	O6qv7 oryzias lat
15	1165.5	84.0	764	O90Y16	O90y16 tetracodon f
16	1147.5	82.7	767	O13133	O13133 oncorhynch
17	754	54.3	163	O9N145	O9n145 macaca mula
18	754	47.8	751	O8BGN0	O8bgn0 xenopus lae
19	654	47.1	712	O6P6Q7	O6p6q7 rattus norv
20	654	47.1	712	AAH62079	Aah62079 rattus no
21	654	47.1	712	O9QXK0	O9qkx0 rattus norv
22	652	47.0	749	O8C497	O8c497 mus musculi
23	651	46.9	712	O99K94	O99k94 mus musculi
24	651	46.9	749	O8C3V4	O8c3v4 mus musculi
25	651	46.9	749	O9D323	O9d323 mus musculi
26	651	46.9	755	O8C8M3	O8c8m3 mus musculi
27	647	46.6	712	AAH35905	Aah35905 homo sapi
28	647	46.6	750	STAX1_HUMAN	P44224 homo sapien
29	646	46.5	757	O764M5	O764m5 sus scrofa
30	646	46.5	757	BAD06318	Bad06318 sus scrofa
31	619	44.6	749	STAX1_MOUSE	P44225 mus musculi

32	606	43.7	754	2	O13131	O13131 oncorhynch
33	602	43.4	718	2	O801Y2	O801y2 carassius a
34	596	42.9	758	2	O90Y17	O90y17 tetracodon f
35	595.5	42.9	528	2	O8JF08	O8jf08 brachydanto
36	595.5	42.9	748	1	STAX4_HUMAN	O14765 homo sapien
37	587.5	42.3	657	2	O8AW24	O8aw24 brachydanto
38	585	42.1	754	2	O13132	O13132 oncorhynch
39	580.5	41.8	749	2	O93598	O93598 brachydanto
40	580.5	41.8	749	2	O6P943	O6p943 brachydanto
41	580.5	41.8	749	2	AAH60933	Aah60933 brachydanto
42	579	41.7	749	1	STAX4_MOUSE	P44228 mus musculi
43	578.5	41.7	651	2	O7Z253	O7z253 brachydanto
44	578.5	41.7	667	2	O8AW20	O8aw20 brachydanto
45	578	41.6	652	2	O7Z277	O7z277 brachydanto

## ALIGNMENTS

RESULT 1	AAH19168	PRELIMINARY;	PRT;	722 AA.
ID	AAH19168			
AC	AAH19168			
DT	02-MAR-2004 (TREMBLrel. 27, Created)			
DT	02-MAR-2004 (TREMBLrel. 27, Last sequence update)			
DT	02-MAR-2004 (TREMBLrel. 27, Last annotation update)			
DE	Signal transducer and activator of transcription 3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STAXIN=FVB/N; TISSUE=Salivary gland;			
RX	MEDLINE=2238257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner U., Shemen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.J.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smallus D.B., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STAXIN=FVB/N; TISSUE=Salivary gland;			
RC	Strausberg R.;			
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC019166; AAH19166.1; --			
SQ	SEQUENCE 722 AA; 83126 MW; 09226A697966D947 CRC64;			
Query Match	99.2%; Score 1377; DB 2; Length 722;			
Best Local Similarity	99.3%; Pred. No. 1,8e-77;			
Matches	269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
QY	1 RCLMEBSRLQTAATAAQQGSGANHPAAVTEKQMLEHILQVRRKVDLEQKRYVE 60			
DB	107 RCLMEBSRLQTAATAAQQGSGANHPAAVTEKQMLEHILQVRRKVDLEQKRYVE 166			
QY	61 NIQDDPENVYTKISQGMOLNNGNSVTRXQMOQLFQMLTALQMRSTYSELAGLLS 120			

```

Db      167 NIODDFENYKTKSCGDMDLNGNNOVTRQMOQLBQMLTALDQMRSSIVSELAGLS 226
Qy      121 AMEYVQKTLDEBLADMKRPEREACIGSPENICLDRLNMTITLAEQSOTRQOIKLSE 180
Db      227 AMEYVQKTLDEBLADMKRROQIACIGSPENICLDRLNMTITLAEQSOTRQOIKLSE 286
Qy      181 LOOKKYSKGDPIYQHRPMLEERIVELPRNLMSAFVVERQPCPMHDPRLVTKGVQFT 240
Db      287 LOOKKYSKGDPIYQHRPMLEERIVELPRNLMSAFVVERQPCPMHDPRLVTKGVQFT 346
Qy      241 TKVRLVKEPELNYOLKIKVCIDKDSGDVAA 271
Db      347 TKVRLVKEPELNYOLKIKVCIDKDSGDVAA 377

RESULT 2
STA3_HUMAN STANDARD; PRT; 770 AA.
AC      P40763; O14916; Q9BWS4;
DT      01-FEB-1995 (Rel. 31, Created)
DT      05-JUL-2004 (Rel. 44, Last sequence update)
DT      01-OCT-2004 (Rel. 45, Last annotation update)
DE      Signal transducer and activator of transcription 3 (Acute-phase
DE      response factor).
GN      Name=STAT3; Synonyms=APRF;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
[1]
RN      SEQUENCE FROM N.A. (ISOFORM 1).
RP      TISSUE=Placenta;
RX      MEDLINE=94208062; PubMed=7512451;
RA      Akita S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA      Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT      "Molecular cloning of APF, a novel IFN-stimulated gene factor 3 p91-
RT      related transcription factor involved in the gp130-mediated signaling
RT      pathway.";
RT      Cell 77:63-71(1994).
[2]
RN      SEQUENCE FROM N.A. (ISOFORM 1).
RP      TISSUE=Placenta;
RX      MEDLINE=9829260; PubMed=9630560;
RA      Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT      "Highly conserved amino-acid sequence between murine STAT3 and a
RT      revised human STAT3 sequence.";
RT      Gene 213:119-124(1998).
[3]
RN      SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT IL6-143.
RP      Rieder M.J., Daniels R.L., da Ponce S.H., Hastings N.C., Ahearn M.O.,
RA      Rajkumar N., Yi Q., Nickerson D.A.;
RT      "SeacLESNPS, NHBI HU6682 program for genomic applications, UW-
RT      PHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RT      Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
[4]
RN      SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RP      TISSUE=Kidney and Pancreas;
RX      MEDLINE=2389257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA      Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
RA      Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Rosh S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettman W., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

```

```

RA      Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [5]
RP      SEQUENCE OF 564-704 FROM N.A.
RC      TISSUE=Liver;
RA      Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RT      Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN      [6]
RP      PHOSPHORYLATION ON SERINE.
RX      MEDLINE=95215843; PubMed=7701321;
RA      Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Xiang S.;
RT      "Requirement of serine phosphorylation for formation of STAT-promoter
RT      complexes.";
RT      Science 267:1990-1994(1995).
CC      -1- FUNCTION: Transcription factor that binds to the interleukin-6
CC      (IL-6)-responsive elements identified in the promoters of various
CC      acute-phase protein genes.
CC      -1- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC      -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC      member (at least STAT1).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC      in response to phosphorylation.
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=1;
CC      IsoId=P40763-1; Sequence=Displayed;
CC      Name=Del-701;
CC      IsoId=P40763-2; Sequence=VSP_010474;
CC      -1- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC      muscle, kidney and pancreas.
CC      -1- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC      LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC      is important for the formation of stable DNA-binding STAT3
CC      homodimers and maximal transcriptional activity.
CC      -1- SIMILARITY: Belongs to the transcription factor STAT family.
CC      -1- SIMILARITY: Contains 1 SH2 domain.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
-----
DR      EMBL; I29277; AAA58374.1; -
DR      EMBL; AJ012463; CAAL0032.1; -
DR      EMBL; AY572796; AAS66986.1; -
DR      EMBL; BC000627; AAH00627.1; -
DR      EMBL; BC014482; AAH14482.1; -
DR      EMBL; AF029311; AAB84254.1; -
DR      PIR; A54444; A54444.
DR      HSSP; P42227; 1BGL.
DR      TRANSFAC; T01493; -
DR      Genew; HGNC:11364; STAT3.
DR      GO; GO:0005737; Cytoplasm; TAS.
DR      GO; GO:0005634; Cytolens; TAS.
DR      GO; GO:0005062; Phenatopoietin/interferon-class (D200-domain. . .); TAS.
DR      GO; GO:0003700; F:transcription factor activity; TAS.
DR      GO; GO:0006928; P:cell motility; TAS.
DR      GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR      GO; GO:0001022; P:negative regulation of transcription from P. . .; TAS.
DR      GO; GO:0007399; P:negative regulation of transcription from P. . .; TAS.
DR      GO; GO:0007165; P:signal transduction; TAS.
DR      InterPro; IPR008567; P53_like_DNA_Bnd.
DR      InterPro; IPR000980; SH2.
DR      InterPro; IPR001217; STAT.
DR      Pfam; PF00017; SH2; 1.
DR      Pfam; PF01017; STAT alpha; 1.
DR      Pfam; PF02864; STAT_Dind; 1.

```

PFam: PRO2865; STAT\_int; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 KW Activator; Alternative splicing; DNA-binding; Nuclear protein;  
 KM Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.  
 FT DOMAIN 580 670 SH2.  
 FT MOD\_RES 705 705 Phosphotyrosine (by JAK) (By similarity).  
 FT MOD\_RES 727 727 Phosphoserine (by similarity).  
 FT VARSPLOC 701 701 Missing (in isoform Del-701).  
 FT FT  
 FT VARIANT 32 32 Q -> K (in isoform Del-701).  
 FT FT  
 FT VARIANT 143 143 M -> I.  
 FT FT  
 FT CONFLICT 288 288 P -> H (in Ref. 1).  
 FT CONFLICT 460 460 Q -> S (in Ref. 1).  
 FT CONFLICT 548 548 K -> N (in Ref. 1).  
 FT CONFLICT 561 561 F -> Y (in Ref. 1).  
 FT CONFLICT 667 667 V -> L (in Ref. 1).  
 FT CONFLICT 730 730 T -> A (in Ref. 1).  
 SQ SEQUENCE 770 AA; 88067 MW; 6C00632211C8012D CRC64;  
 Query Match 99.2%; Score 1377; DB 1; Length 770;  
 Best Local Similarity 99.3%; Pred. No. 2e-77;  
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RCLWESSRLQTAATAAQQGGQANHPAAVTEKQOMLEQHLQDVKKVQDLEQKMYVE 60  
 DB 107 RCLWESSRLQTAATAAQQGGQANHPAAVTEKQOMLEQHLQDVKKVQDLEQKMYVE 166  
 OY 61 NLQDQFDNNYKTKSQGDMQDNGNNSVTRQKMOQLQMLTALQMRRSYSELGILLS 120  
 DB 167 NLQDQFDNNYKTKSQGDMQDNGNNSVTRQKMOQLQMLTALQMRRSYSELGILLS 226  
 OY 121 AMEYVQKLTBELEADMKRRPEIACIGSPNCLDRLNMTSLAESQLTROQIKKLEB 180  
 DB 227 AMEYVQKLTBELEADMKRRQIACIGSPNCLDRLNMTSLAESQLTROQIKKLEB 286  
 OY 181 LQQRVSYKGDPIVQHPMLFEERIVLFLNLMKSAFVERQPCMPHMDPRPLVIKTGVQFT 240  
 DB 287 LQQRVSYKGDPIVQHPMLFEERIVLFLNLMKSAFVERQPCMPHMDPRPLVIKTGVQFT 346  
 OY 241 TKVRLVKEPELNLQTKIKVCIDKDSGVAA 271  
 DB 347 TKVRLVKEPELNLQTKIKVCIDKDSGVAA 377

RESULT 3  
 ID STAS\_MOUSE STANDARD; PRT; 770 AA.  
 AC P42227;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Signal transducer and activator of transcription 3 (Acute-phase response factor).  
 GN Name=Stat3; Synonyms=Apf1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_Taxid=10090;  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185 AND 632-640.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=94208062; PubMed=7512451;  
 RA Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T., Yoshida K., Sudo T., Naruto M., Kishimoto T.;  
 RT "Molecular cloning of APF1, a novel IFN-gamma-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway.";  
 RL Cell 77:63-71(1994).  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).

RC TISSUE=Thymus;  
 RX MEDLINE=94188718; PubMed=8140422;  
 RA Zhong Z., Wen Z., Darnell J.E. Jr.;  
 RT "Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6.";  
 RL Science 264:95-98(1994).  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).  
 RC TISSUE=Brain;  
 RX MEDLINE=95014185; PubMed=7523373;  
 RA Raz R., Durbin J.E., Levy D.E.;  
 RT "Acute phase response factor and additional members of the interferon-stimulated gene factor 3 family integrate diverse signals from cytokines, interferons, and growth factors.";  
 RL J. Biol. Chem. 269:24391-24395(1994).  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM STAT3B).  
 RC STRAIN=BALB/c; and C57BL/6; TISSUE=Liver;  
 RX MEDLINE=96016116; PubMed=7568080;  
 RA Schaefer T.S., Sanders L.K., Nathans D.;  
 RT "Cooperative transcriptional activity of Jun and Stat3 beta, a short form of Stat3.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).  
 RC STRAIN=129/SvJ;  
 RX PubMed=11161808; DOI=10.1006/geno.2000.6433;  
 RA Miyoshi K., Cui Y., Riedlinger G., Robinson P., Lehoczy J., Zon L., Oka T., Dewar K., Hennighausen L.;  
 RT "Structure of the mouse Stat 3/5 locus: evolution from Drosophila to zebrafish to mouse.";  
 RL Genomics 71:150-155(2001).  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).  
 RC STRAIN=C57BL/6J; and NOD/LtJ;  
 RA Davoodi-Semlomi A., She U.-X.;  
 RT "A mutant Stat5b with weaker DNA binding defines a key defective pathway in non-obese diabetic (NOD) mice.";  
 RL Submitted (May-2003) to the EMBL/Genbank/DBJ databases.  
 RN  
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).  
 RC STRAIN=FVB/N; TISSUE=Mammary gland;  
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN  
 RP PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.  
 RX MEDLINE=95354205; PubMed=7543024;  
 RA Wen Z., Zhong Z., Darnell J.E. Jr.;  
 RT "Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation.";  
 RL Cell 82:241-250(1995).  
 RN  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.  
 RX MEDLINE=98334373; PubMed=9671298;

RA Becker S., Groner B., Mueller C.W.;  
 RT "Three-dimensional structure of the Stat3beta homodimer bound to  
 RT DNA.";  
 RL Nature 394:145-151 (1998).  
 CC -1- FUNCTION: Transcription factor that binds to the interleukin-6  
 CC (IL-6)-responsive elements identified in the promoters of various  
 CC acute-phase protein genes. STAT3B interacts with the N-terminal  
 CC part of JUN to activate such promoters in a cooperative way.  
 CC -1- PATHWAY: Involved in the gp130-mediated signaling pathway.  
 CC -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family  
 CC member (at least STAT1). (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus  
 CC in response to phosphorylation.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=Stat3A;  
 CC IsoId=P42227-1; Sequence=Displayed;  
 CC Name=Stat3B;  
 CC IsoId=P42227-2; Sequence=VSP\_006287;  
 CC Name=Del-701;  
 CC IsoId=P42227-3; Sequence=VSP\_010475;  
 CC -1- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and  
 CC kidney. STAT3B is also detected in the liver, although in a much  
 CC less abundant manner.  
 CC -1- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,  
 CC LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation  
 CC is important for the formation of stable DNA-binding STAT3  
 CC homodimers and maximal transcriptional activity (By similarity).  
 CC -1- SIMILARITY: Belongs to the transcription factor STAT family.  
 CC -1- SIMILARITY: Contains 1 SH2 domain.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; L29278; AAA37254.1; -;  
 DR EMBL; U06922; AAA19452.1; -;  
 DR EMBL; U08378; AAA56668.1; -;  
 DR EMBL; U30709; AAC52612.1; -;  
 DR EMBL; AF246978; AAL59017.1; -;  
 DR EMBL; AY299489; AAQ75418.1; -;  
 DR EMBL; AY299490; AAQ75419.1; -;  
 DR EMBL; BC003806; AA03806.1; -;  
 DR PIR; I49508; I49508.  
 DR PDB; 1BG1; X-ray; A=1-722.  
 DR TRANSFAC; T01574; -;  
 DR MGD; MGI:103038; Stat3.  
 DR GO; GO:0005737; Cytoplasm; IDA.  
 DR GO; GO:0005634; Cnucleus; IDA.  
 DR GO; GO:0005886; Cplasma membrane; IDA.  
 DR GO; GO:0003677; F-DNA binding; IDA.  
 DR GO; GO:0005515; F-protein binding; IPI.  
 DR GO; GO:0015653; F-transcriptional activator activity; IDA.  
 DR GO; GO:0006357; P-regulation of transcription from Pol II pro. .; IDA.  
 DR GO; GO:0006357; P-regulation of transcription from Pol II pro. .; IDA.  
 DR InterPro; IPR008967; P52\_like\_DNA\_bnd.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF01017; STAT\_alpha; 1.  
 DR Pfam; PF02864; STAT\_bind; 1.  
 DR Pfam; PF02865; STAT\_int; 1.  
 DR PROSITE; PSS0001; SH2; 1.  
 DR 3D-structure; Activator; Acute phase; Alternative splicing;  
 DR Direct protein sequencing; DNA-binding; Nuclear protein;  
 DR Phosphorylation; SH2 domain; Transcription regulation.  
 KW DOMAIN 580 670  
 FT MOD\_RES 705 705 Phosphotyrosine (by JAK) (By similarity).  
 FT MOD\_RES 727 727 Phosphoserine.

FT	VARSPLIC	716	770	TTCSNTIDLPMSPRTLSIMQFGNNGEAPSGQFESLT FMDLTSSECATSPM -> FIDAVMK (in isoform Stat3B). /FTid=VSP_006287. Missing (in isoform Del-701). S->A: Decreased transcriptional activation. B -> K (in Ref. 2). S -> T (in Ref. 2 and 4). M -> I (in Ref. 1).
FT	VARSPLIC	701	701	
FT	MUTAGEN	727	727	
FT	CONFLICT	16	16	
FT	CONFLICT	25	25	
FT	CONFLICT	394	394	
FT	HELI	139	180	
FT	TURN	181	182	
FT	TURN	197	198	
FT	HELI	199	237	
FT	TURN	238	237	
FT	HELI	239	251	
FT	TURN	252	253	
FT	HELI	261	280	
FT	TURN	294	295	
FT	TURN	297	301	
FT	HELI	302	320	
FT	STRAND	321	328	
FT	TURN	330	331	
FT	TURN	333	334	
FT	TURN	336	337	
FT	STRAND	338	340	
FT	TURN	341	342	
FT	STRAND	345	351	

Query Match 99.2%; Score 1377; DB 1; Length 770;  
 Best Local Similarity 99.3%; Pred. No. 2e-77; 1; Indels 0; Gaps 0;

Matches 269; Conservative 1; Mismatches 1;

QY	1	RCIMESRLLOPAATPAAGGQANPTAAVTEKQOMLEQHIQDYKRYQDLEQKKAYE 60
DB	107	RCIMESRLLOPAATPAAGGQANPTAAVTEKQOMLEQHIQDYKRYQDLEQKKAYE 166
QY	61	NIQDDPFENYKTLKSGDMQDNGNNSVTRQMOQLEQMTALDQRRSIVSELAGLS 120
DB	167	NIQDDPFENYKTLKSGDMQDNGNNSVTRQMOQLEQMTALDQRRSIVSELAGLS 226
QY	121	AMEYQKLTDEBLADMKRRPFIACIGPPNICLRLEWITSLSQOTROQIKKEE 180
DB	227	AMEYQKLTDEBLADMKRRPFIACIGPPNICLRLEWITSLSQOTROQIKKEE 286
QY	181	LOQKSYKQDPVQHRPMLERIVELFRNLMSAFVERQPCMPHDPDELVIKTVGQFT 240
DB	287	LOQKSYKQDPVQHRPMLERIVELFRNLMSAFVERQPCMPHDPDELVIKTVGQFT 346
QY	241	TVRRLIVKPELNIQIKVICIDKSGDYAA 271
DB	347	TVRRLIVKPELNIQIKVICIDKSGDYAA 377

RESULT 4  
 STD3 RAT  
 ID STD3 RAT STANDARD; PRT; 770 AA.  
 AC P52631;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Signal transducer and activator of transcription 3.  
 GN Name=Stat3;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 ON NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=96102059; PubMed=8530402;  
 RA Ripberger J.A., Filtz S., Richter K., Hooke G.M., Lottepeich F.,

RA Fey G.H.;  
 RT "Transcription factors Stat3 and Stat5b are present in rat liver  
 RT nuclei late in an acute phase response and bind interleukin-6 response  
 RT elements";  
 RL J. Biol. Chem. 270:29998-30006(1995).  
 CC -1- FUNCTION: Transcription factor that binds to the interleukin-6  
 CC (IL-6)-responsive elements identified in the promoters of various  
 CC acute-phase protein genes.  
 CC -1- PATHWAY: Involved in the gp130-mediated signaling pathway.  
 CC -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family  
 CC member (at least STAT1) (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus  
 CC in response to phosphorylation (By similarity).  
 CC -1- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,  
 CC ILF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation  
 CC is important for the formation of stable DNA-binding STAT3  
 CC homodimers and maximal transcriptional activity (By similarity).  
 CC -1- SIMILARITY: Belongs to the transcriptional factor STAT family.  
 CC -1- SIMILARITY: Contains 1 SH2 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X91810; CAA62920.1; -.  
 DR HSSP: P42227; 1BG1.  
 DR RSD; 3772; Stat3.  
 DR InterPro: IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro: IPR000980; SH2.  
 DR InterPro: IPR001217; STAT.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF01017; STAT\_alpha; 1.  
 DR Pfam: PF02864; STAT\_bind; 1.  
 DR Pfam: PF02865; STAT\_int; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR PROSITE: PS0001; SH2; 1.  
 KW Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;  
 KW Transcription regulation.  
 FT DOMAIN 580 670  
 FT MOD\_RES 705 705 Phosphotyrosine (by JAK) (By similarity).  
 FT MOD\_RES 727 727 Phosphoserine (By similarity).  
 SQ SEQUENCE 770 AA; 88039 MW; D74A0C76954754ED CRC64;

Query Match 99.0%; Score 1374; DB 1; Length 770;  
 Best Local Similarity 98.9%; Pred. No. 3e-77;  
 Matches 268; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RCLMBESRLIQTATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 60  
 DB 107 RCLMBESRLIQTATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 166  
 OY 61 NLADDFPNYKTLKSOGDMODLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 120  
 DB 167 NLADDFPNYKTLKSOGDMODLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 226  
 OY 121 AMEYVQKTLTDEELADWKRREPIACIGPPNICDRLNNWITSLAESOLQTRQOIKKLEE 180  
 DB 227 AMEYVQKTLTDEELADWKRREPIACIGPPNICDRLNNWITSLAESOLQTRQOIKKLEE 286  
 OY 181 LQOKVSYKGDPIVQHRPMLERIVLEFNNLMKSAFVVERQPCMPHPRPIVITKGVQFT 240  
 DB 287 LQOKVSYKGDPIVQHRPMLERIVLEFNNLMKSAFVVERQPCMPHPRPIVITKGVQFT 346  
 OY 241 TKVRLIVKFPPELNTQKIKVCIDKDSGDVAA 271  
 DB 347 TKVRLIVKFPPELNTQKIKVCIDKDSGDVAA 377

RESULT 5

STAT3 BOVIN  
 ID STAT3 BOVIN STANDARD; PRT; 770 AA.  
 AC P61635;  
 DT 05-JUN-2004 (Rel. 44, Created)  
 DT 05-JUN-2004 (Rel. 44, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DB Signal transducer and activator of transcription 3.  
 GN Name=STAT3;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Seyfert H.M., Wheeler T.T., Woolenaar A., Pitra C.;  
 RT "The STAT3B-encoding gene was flipped across the STAT3/STAT5A-locus  
 RT during ruminant evolution.";  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Transcription factor that binds to the interleukin-6  
 CC (IL-6)-responsive elements identified in the promoters of various  
 CC acute-phase protein genes (By similarity).  
 CC -1- PATHWAY: Involved in the gp130-mediated signaling pathway.  
 CC -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family  
 CC member (at least STAT1) (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus  
 CC in response to phosphorylation (By similarity).  
 CC -1- SIMILARITY: Belongs to the transcription factor STAT family.  
 CC -1- SIMILARITY: Contains 1 SH2 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF620655; CAF06182.1; -.  
 DR PROSITE: PS0001; SH2; 1.  
 KW Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;  
 KW Transcription regulation.  
 FT DOMAIN 580 670  
 FT MOD\_RES 705 705 Phosphotyrosine (by JAK) (By similarity).  
 FT MOD\_RES 727 727 Phosphoserine (By similarity).  
 SQ SEQUENCE 770 AA; 87974 MW; 9CEB147C73B83274 CRC64;

Query Match 98.7%; Score 1370; DB 1; Length 770;  
 Best Local Similarity 98.9%; Pred. No. 5.3e-77;  
 Matches 268; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RCLMBESRLIQTATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 60  
 DB 107 RCLMBESRLIQTATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 166  
 OY 61 NLADDFPNYKTLKSOGDMODLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 120  
 DB 167 NLADDFPNYKTLKSOGDMODLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 226  
 OY 121 AMEYVQKTLTDEELADWKRREPIACIGPPNICDRLNNWITSLAESOLQTRQOIKKLEE 180  
 DB 227 AMEYVQKTLTDEELADWKRREPIACIGPPNICDRLNNWITSLAESOLQTRQOIKKLEE 286  
 OY 181 LQOKVSYKGDPIVQHRPMLERIVLEFNNLMKSAFVVERQPCMPHPRPIVITKGVQFT 240  
 DB 287 LQOKVSYKGDPIVQHRPMLERIVLEFNNLMKSAFVVERQPCMPHPRPIVITKGVQFT 346  
 OY 241 TKVRLIVKFPPELNTQKIKVCIDKDSGDVAA 271  
 DB 347 TKVRLIVKFPPELNTQKIKVCIDKDSGDVAA 377



```

RESULT 6
Q6DV79 PRELIMINARY; PRT; 771 AA.
AC Q6DV79;
DT 01-OCT-2004 (TREMBlrel. 28, Created)
DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Signal transducer and activator of transcription 3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou G.Y., Leung F.C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY641397; AAT64887.1; -
SQ SEQUENCE 771 AA; 88174 MW; 71AC855C5DEC03E2 CRC64;

Query Match 97.4%; Score 1352; DB 2; Length 771;
Best Local Similarity 97.4%; Pred. No. 7e-76;
Matches 264; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RCLMEESRLQTAATTAAGGGAHPTAAVTEKQMLEQHIQDVRAKRVQDLEQKKKYE 60
DB 107 RCLMEESRLQTAATTAAGGGAHPTAAVTEKQMLEQHIQDVRAKRVQDLEQKKKYE 166
QY 61 NLQDDPDPFNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRSSIVSEIAGLIS 120
DB 167 NLQDDPDPFNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRSSIVSEIAGLIS 226
QY 121 AMEYVQKTLTDEBLADWKRREPIACIGPPNICDLRLNWTSLAESQIQRIKKLEE 180
DB 227 AMEYVQKTLTDEBLADWKRREPIACIGPPNICDLRLNWTSLAESQIQRIKKLEE 286
QY 121 AMEYVQKTLTDEBLADWKRREPIACIGPPNICDLRLNWTSLAESQIQRIKKLEE 180
DB 227 AMEYVQKTLTDEBLADWKRREPIACIGPPNICDLRLNWTSLAESQIQRIKKLEE 286
QY 181 LQOKSVYKGDPIVQHRPMLSEIRIVELFRLNMSAFVVERQPCMPHPRPIVTKGVQFT 240
DB 287 LQOKSVYKGDPIVQHRPMLSEIRIVELFRLNMSAFVVERQPCMPHPRPIVTKGVQFT 346
QY 241 TKVRLVVKPELNYQIKIKVCIDKSGDVAA 271
DB 347 TKVRLVVKPELNYQIKIKVCIDKSGDVAA 377

RESULT 7
Q9PVX8 PRELIMINARY; PRT; 769 AA.
AC Q9PVX8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Stat 3.
GN Name=stat 3;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20107399; PubMed=10642787;
RA Nishinakamura R., Matsuno Y., Matsuda T., Arizumi T., Heike T.,
RA Asahina M., Yokota T.;
RL "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT embryos independent of BMP-4."
RL Dev. Biol. 216:481-490 (1999).
DR EMBL; AB017701; BAA6061.1; -
DR HSRF; P42227; IBG1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.

```

```

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PR00017; SH2; 1.
DR Pfam; PR01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 769 AA; 87974 MW; 0905C03263303069 CRC64;

Query Match 93.9%; Score 1304; DB 2; Length 769;
Best Local Similarity 93.0%; Pred. No. 6.8e-73;
Matches 252; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 RCLMEESRLQTAATTAAGGGAHPTAAVTEKQMLEQHIQDVRAKRVQDLEQKKKYE 60
DB 107 RCLMEESRLQTAATTAAGGGAHPTAAVTEKQMLEQHIQDVRAKRVQDLEQKKKYE 166
QY 61 NLQDDPDPFNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRSSIVSEIAGLIS 120
DB 167 NLQDDPDPFNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRSSIVSEIAGLIS 226
QY 121 AMEYVQKTLTDEBLADWKRREPIACIGPPNICDLRLNWTSLAESQIQRIKKLEE 180
DB 227 AMEYVQKTLTDEBLADWKRREPIACIGPPNICDLRLNWTSLAESQIQRIKKLEE 286
QY 181 LQOKSVYKGDPIVQHRPMLSEIRIVELFRLNMSAFVVERQPCMPHPRPIVTKGVQFT 240
DB 287 LQOKSVYKGDPIVQHRPMLSEIRIVELFRLNMSAFVVERQPCMPHPRPIVTKGVQFT 346
QY 241 TKVRLVVKPELNYQIKIKVCIDKSGDVAA 271
DB 347 NKVRLVVKPELNYQIKIKVCIDKSGDVAA 377

RESULT 8
Q7ZXK3 PRELIMINARY; PRT; 766 AA.
AC Q7ZXK3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Stat3-A protein.
GN Name=stat3-A;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Embryo;
RA MEDLINE=22388257; PubMed=12477932;
RA Strassberg R.V., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abtannon R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RL "Generation and initial analysis of more than 15,000 full-length human
RT

```

RT and mouse cDNA sequence: "  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative";  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; BC044717; AAA44717.1; -;  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF01017; STAT\_alpha; 1.  
 DR Pfam; PF02864; STAT\_bird; 1.  
 DR Pfam; PF02865; STAT\_int; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 SQ SEQUENCE 766 AA; 87599 MW; 31018A3321CCEB9C CRC64;

Query Match 92.7%; Score 1286; DB 2; Length 766;  
 Best Local Similarity 91.1%; Pred. No. 8.9e-72;  
 Matches 247; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 1 RCLMEESRLTQTATAAOCGQANHTAAVTEKQOMLEOHLOVDRKRVODLEQKMYVE 60  
 DB 107 RCLMEESRLTQTATAAOCGQANHTAAVTEKQOMLEOHLOVDRKRVODLEQKMYVE 166  
 QY 61 NLQDDPFDNYKTLKSGDMODLNGNNSVTRQKMOLEOMLTALDQMRRSIVSLAGL 120  
 DB 167 NLQDDPFDNYKTLKSGDLSLNGNNSVTRQKMOLEOMLTALDQMRRSIVSLAGL 226  
 QY 121 AMEVYQKTLTDEBLADMKRRPEIACIGGPNICLDRLNWTITSLAESQLOTRQOIRKLE 180  
 DB 227 AMEVYQKTLTDEBLADMKRRPEIACIGGPNICLDRLNWTITSLAESQLOTRQOIRKLE 286  
 QY 181 LQOKVSYKGDPIVQHRPMLERIVELFRNLKSAFVEROPCPMPHDPRLVIKTVQ 240  
 DB 287 LQOKVSYKGDPIVQHRPMLERIVELFRNLKSAFVEROPCPMPHDPRLVIKTVQ 346  
 QY 241 TKVRLVLPPELNYQKIKVICIDKSGDVAA 271  
 DB 347 NKVRLVLPPELNYQKIKVICIDKSGDVAA 377  
 RESULT 9  
 Q7ZTS5 PRELIMINARY; PRT; 414 AA.  
 AC Q7ZTS5;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DR Stat3 protein.  
 GN Name=stat3;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Actinoptera; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OK NCBI\_taxid=7955;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Whole body;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Hellon E., Kettleman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalski U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Whole body;  
 RA Strausberg R.;  
 RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; BC045276; AAA5276.1; -;  
 DR ZFIN; ZDB-GENE-980526-66; stat3.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF01017; STAT\_alpha; 1.  
 DR Pfam; PF02864; STAT\_bird; 1.  
 DR Pfam; PF02865; STAT\_int; 1.  
 SQ SEQUENCE 414 AA; 48253 MW; 0FFD1B509B7526BD CRC64;

Query Match 85.8%; Score 1191.5; DB 2; Length 414;  
 Best Local Similarity 84.6%; Pred. No. 3.3e-66;  
 Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;

QY 1 RCLMEESRLTQTATAAOCGQANHTAAVTEKQOMLEOHLOVDRKRVODLEQKMYVE 60  
 DB 107 RCLMEESRLTQT-ATTAQDDQVAHPITGVTEKQOILEHNLQDTRKRVODLEQKMYVE 165  
 QY 61 NLQDDPFDNYKTLKSGDM-ODLNGNNO-SVTRQKMOLEOMLTALDQMRRSIVSLAGL 118  
 DB 166 NLQDDPFDNYKTLKSGDSLNGNNSQAATROGMSQLEOMLSLDDQRLRQIVTEMAGL 225  
 QY 119 LSAEYVQKTLTDEBLADMKRRPEIACIGGPNICLDRLNWTITSLAESQLOTRQOIRKLE 178  
 DB 226 LSAEYVQKTLTDEBLADMKRRQIACIGGPNICLDRLNWTITSLAESQLOTRQOIRKLE 285  
 QY 179 ELQOKVSYKGDPIVQHRPMLERIVELFRNLKSAFVEROPCPMPHDPRLVIKTVQ 238  
 DB 286 ELQOKVSYKGDPIVQHRPALKEKIVDLFRNLKSAFVEROPCPMPHDPRLVIKTVQ 345  
 QY 239 FTTKRLVLPPELNYQKIKVICIDKSGDVAA 271  
 DB 346 FTTKRLVLPPELNYQKIKVICIDKSGDVAA 378

RESULT 10  
 Q6NY46 PRELIMINARY; PRT; 786 AA.  
 AC Q6NY46;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)  
 DE Stat3 protein.  
 GN Name=stat3;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.W.,  
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,  
 Jones S.J., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Kidney;  
 RC Strausberg R.;  
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC068320; AAH68320.1; -  
 DR InterPro: IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro: IPR000980; SH2.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF01017; STAT\_alpha; 1.  
 DR Pfam: PF02864; STAT\_bind; 1.  
 DR Pfam: PF02865; STAT\_int; 1.  
 DR PROSITE: P55001; SH2; 1.  
 SQ SEQUENCE 786 AA; 90039 MW; FC7371D0B0E5447E CRC64;  
 Query Match 85.8%; Score 1191.5; DB 2; Length 786;  
 Best Local Similarity 84.6%; Pred. No. 6,8e-66;  
 Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;  
 QY 1 RCLMEBSRLLOTPAATPAOQGGQANHPPTAAVTEKQOMLEQHLQDVKKRVODLEQKKKVE 60  
 DB 107 RCLMEBSRLLOTPAATPAOQGGQANHPPTAAVTEKQOMLEQHLQDVKKRVODLEQKKKVE 165  
 QY 61 NLODPDEFNYKYLKSGGDM-ODLNGNNO-SVTRQKQOOLEMTLTALDQKRSTVSISLAL 118  
 DB 166 NLODPDEFNYKYLKSGGDM-ODLNGNNO-SVTRQKQOOLEMTLTALDQKRSTVSISLAL 225  
 QY 119 ISAMEVQKTLTDEBLADWKRREPELACIGGPNICIDRLNNWITSLSAESQLOTRQOIKTL 178  
 DB 226 ISAMFVQGNLTDEBLADWKRREPELACIGGPNICIDRLNNWITSLSAESQLOTRQOIKTL 285  
 QY 179 BELQOKVSYKGPPIVQHRPMLERIVELFRNLKSAFVVERQPCPMHEDRPLVITKGVQ 238  
 DB 286 BELQOKVSYKGPPIVQHRPMLERIVELFRNLKSAFVVERQPCPMHEDRPLVITKGVQ 345  
 QY 239 FTTKRLVLKPEPLNYOLKIKVICIDKSGDVAA 271  
 DB 346 FTTKRLVLKPEPLNYOLKIKVICIDKSGDVAA 378

RESULT 11  
 ID AAH68320 PRELIMINARY; PRT; 786 AA.  
 AC AAH68320;  
 DT 10-MAY-2004 (TrEMBLrel. 27, Created)  
 DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Stat3 protein.  
 GN STAT3.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.W.,  
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,  
 Jones S.J., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Kidney;  
 RC Strausberg R.;  
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC068320; AAH68320.1; -  
 DR InterPro: IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro: IPR000980; SH2.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF01017; STAT\_alpha; 1.  
 DR Pfam: PF02864; STAT\_bind; 1.  
 DR Pfam: PF02865; STAT\_int; 1.  
 DR PROSITE: P55001; SH2; 1.  
 SQ SEQUENCE 786 AA; 90039 MW; FC7371D0B0E5447E CRC64;  
 Query Match 85.8%; Score 1191.5; DB 2; Length 786;  
 Best Local Similarity 84.6%; Pred. No. 6,8e-66;  
 Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;  
 QY 1 RCLMEBSRLLOTPAATPAOQGGQANHPPTAAVTEKQOMLEQHLQDVKKRVODLEQKKKVE 60  
 DB 107 RCLMEBSRLLOTPAATPAOQGGQANHPPTAAVTEKQOMLEQHLQDVKKRVODLEQKKKVE 165  
 QY 61 NLODPDEFNYKYLKSGGDM-ODLNGNNO-SVTRQKQOOLEMTLTALDQKRSTVSISLAL 118  
 DB 166 NLODPDEFNYKYLKSGGDM-ODLNGNNO-SVTRQKQOOLEMTLTALDQKRSTVSISLAL 225  
 QY 119 ISAMEVQKTLTDEBLADWKRREPELACIGGPNICIDRLNNWITSLSAESQLOTRQOIKTL 178  
 DB 226 ISAMFVQGNLTDEBLADWKRREPELACIGGPNICIDRLNNWITSLSAESQLOTRQOIKTL 285  
 QY 179 BELQOKVSYKGPPIVQHRPMLERIVELFRNLKSAFVVERQPCPMHEDRPLVITKGVQ 238  
 DB 286 BELQOKVSYKGPPIVQHRPMLERIVELFRNLKSAFVVERQPCPMHEDRPLVITKGVQ 345  
 QY 239 FTTKRLVLKPEPLNYOLKIKVICIDKSGDVAA 271  
 DB 346 FTTKRLVLKPEPLNYOLKIKVICIDKSGDVAA 378

RESULT 12  
 093599

ID 093599 PRELIMINARY; PRT; 806 AA.  
 AC 093599;  
 DT 01-NOV-1998 (Tremblrel. 03, Created)  
 DT 01-NOV-1998 (Tremblrel. 03, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE Transcription factor.  
 GN Name=atac3;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NC NCBITaxID=7955;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Oates A.C.;  
 RL Thesis (1998), University of Melbourne, Australia.  
 DR EMBL; AJ005693; CAA06677.1; -.  
 DR HSSP; P42227; 18Q1.  
 DR ZFIN; ZDB-GENE-980526-68; stat3.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR009807; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF01017; STAT\_alpha; 2.  
 DR Pfam; PF02864; STAT\_bind; 1.  
 DR Pfam; PF02865; STAT\_int; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 SQ SEQUENCE 806 AA; 92151 MW; 74BC4EA401C3C942 CRC64;

Query Match 85.8%; Score 1191.5; DB 2; Length 806;  
 Best Local Similarity 84.6%; Pred. NO. 7e-66;  
 Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;

QY 1 RCLWESRLIQTATAAQQGGQANHPTRAAYVTEKQMLEQHLOVRRKVDLEQMKVVE 60  
 DB 107 RCLWESRLIQT-ATTSQDSGAHPGTGVTEKQMLEHLOVRRKVDLEQMKVVE 165  
 QY 61 NLQDDPFPNFKYTKLSQSGDM-ODLNGNNO-SVTRQKMOQLEOMLTALDQMRISYSELAGL 118  
 DB 166 NLQDDPFPNFKYTKLSQSGDM-ODLNGNNO-SVTRQKMOQLEOMLTALDQMRISYSELAGL 225  
 QY 119 LSAMEYQKTLTDELDADWKRPRPIACIGGPPNICDLRLNWITSLSAQLOTRQOIKKL 178  
 DB 226 LSAMEYQKTLTDELDADWKRPRPIACIGGPPNICDLRLNWITSLSAQLOTRQOIKKL 285  
 QY 179 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVEROPCPMHDPDPVYIKTVQ 238  
 DB 286 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVEROPCPMHDPDPVYIKTVQ 345  
 QY 239 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 271  
 DB 346 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 378

## RESULT 13

Q6DVFP3 PRELIMINARY; PRT; 765 AA.  
 AC Q6DVFP3;  
 DT 01-OCT-2004 (Tremblrel. 28, Created)  
 DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)  
 DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)  
 DE Signal transducer and activator of transcription factor 3.  
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorphi; Acanthopterygii; Percomorphi; Atherinomorpha;  
 OC Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.

Query Match 85.2%; Score 1182.5; DB 2; Length 765;  
 Best Local Similarity 83.9%; Pred. No. 2.4e-65;  
 Matches 229; Conservative 22; Mismatches 19; Indels 3; Gaps 3;

OX NCBITaxID=8090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liu R., Hong Y.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF639947; AAT64912.1; -.  
 SQ SEQUENCE 765 AA; 87566 MW; F5D01A08748BC703 CRC64;

Query Match 85.2%; Score 1182.5; DB 2; Length 765;  
 Best Local Similarity 83.9%; Pred. No. 2.4e-65;  
 Matches 229; Conservative 22; Mismatches 19; Indels 3; Gaps 3;

QY 1 RCLWESRLIQTATAAQQGGQANHPTRAAYVTEKQMLEQHLOVRRKVDLEQMKVVE 60  
 DB 107 RCLWESRLIQT-ATTSQDSGAHPGTGVTEKQMLEHLOVRRKVDLEQMKVVE 165  
 QY 61 NLQDDPFPNFKYTKLSQSGDM-ODLNGNNO-SVTRQKMOQLEOMLTALDQMRISYSELAGL 118  
 DB 166 NLQDDPFPNFKYTKLSQSGDM-ODLNGNNO-SVTRQKMOQLEOMLTALDQMRISYSELAGL 225  
 QY 119 LSAMEYQKTLTDELDADWKRPRPIACIGGPPNICDLRLNWITSLSAQLOTRQOIKKL 178  
 DB 226 LSAMEYQKTLTDELDADWKRPRPIACIGGPPNICDLRLNWITSLSAQLOTRQOIKKL 285  
 QY 179 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVEROPCPMHDPDPVYIKTVQ 238  
 DB 286 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVEROPCPMHDPDPVYIKTVQ 345  
 QY 239 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 271  
 DB 346 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 378

## RESULT 14

Q6GUE7 PRELIMINARY; PRT; 765 AA.  
 AC Q6GUE7;  
 DT 05-JUN-2004 (Tremblrel. 27, Created)  
 DT 05-JUN-2004 (Tremblrel. 27, Last sequence update)  
 DT 05-JUN-2004 (Tremblrel. 27, Last annotation update)  
 DE Signal transducer and activator of transcription 3 isoform 1.  
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorphi; Acanthopterygii; Percomorphi; Atherinomorpha;  
 OC Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.  
 NC NCBITaxID=8090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Liu R., Hong Y.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY641434; AAT6364.1; -.  
 DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF01017; STAT\_alpha; 1.  
 DR Pfam; PF02864; STAT\_bind; 1.  
 DR Pfam; PF02865; STAT\_int; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 SQ SEQUENCE 765 AA; 89643 MW; 81F231BDE27DE938 CRC64;

Query Match 85.2%; Score 1182.5; DB 2; Length 765;  
 Best Local Similarity 83.9%; Pred. No. 2.4e-65;  
 Matches 229; Conservative 22; Mismatches 19; Indels 3; Gaps 3;

QY 1 RCLWESRLIQTATAAQQGGQANHPTRAAYVTEKQMLEQHLOVRRKVDLEQMKVVE 60  
 DB 107 RCLWESRLIQT-ATTSQDSGAHPGTGVTEKQMLEHLOVRRKVDLEQMKVVE 165  
 QY 61 NLQDDPFPNFKYTKLSQSGDM-ODLNGNNO-SVTRQKMOQLEOMLTALDQMRISYSELAGL 118  
 DB 166 NLQDDPFPNFKYTKLSQSGDM-ODLNGNNO-SVTRQKMOQLEOMLTALDQMRISYSELAGL 225

```

QY 119 ISAMEYVQTLTDEDEIADWKREPELACIGPPNICDLRLNMTTSLAESQLQTRQIKKL 178
DB 226 LTAMDYVQNLTLDEELADWKRRQOIAICIGPPNICDLRLNMTTSLAESQLQTRQIKKL 285
QY 179 ERLQOKVSKGDPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 238
DB 286 ERLQOKVSKGDPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 345
QY 239 FTTKVRLLVKKPELNYOLKRVICIDKDSGDVAA 271
DB 346 FTKKVRLLVKKPELNYOLKRVICIDKDSGDVAA 378

```

## RESULT 15

```

Q90Y16 PRELIMINARY; PRT; 764 AA.
ID Q90Y16;
AC Q90Y16;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE STAT3.
OS Tetraodon fluviatilis (Puffer fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_Taxid=47145;
RN [1]
RP SEQUENCE FROM N.A.
RA Sung S.C., Fan T.J., Liu J.H., Chou C.M., Huang C.J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF307106; AAL09415.1; -.
DR HSSP; P42227; 1BG1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_CNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR01217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SQ SEQUENCE 764 AA; 87411 MW; E661FPE18BEPDBE CRC64;

```

```

Query Match 84.0%; Score 1165.5; DB 2; Length 764;
Best Local Similarity 83.2%; Pred.No. 2.7e-64;
Matches 227; Conservative 24; Mismatches 19; Indels 3; Gaps 3;

```

```

QY 1 RCIMESRLLQTAATAAQGGQANHPATAVTEKQOMLEQHLQDVRKRYVDLEQKKMYVE 60
DB 107 RCIMESRRLQTAATAAQGGQANHPATAVTEKQOMLEQHLQDVRKRYVDLEQKKMYVE 165
QY 61 NIQDDFDFNYKTLKSGQDM-QDLNGNQ-SVTRQKXQOLEOMLTALDQWRSIVSELGL 118
DB 166 NIQDDFDFNYKTLKSGQDM-QDLNGNQ-SVTRQKXQOLEOMLTALDQWRSIVSELGL 225
QY 119 ISAMEYVQTLTDEDEIADWKREPELACIGPPNICDLRLNMTTSLAESQLQTRQIKKL 178
DB 226 LTAMDYVQNLTLDEELADWKRRQOIAICIGPPNICDLRLNMTTSLAESQLQTRQIKKL 285
QY 179 ERLQOKVSKGDPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 238
DB 286 ERLQOKVSKGDPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 345
QY 239 FTTKVRLLVKKPELNYOLKRVICIDKDSGDVAA 271
DB 346 FTKKVRLLVKKPELNYOLKRVICIDKDSGDVAA 378

```

Search completed: December 2, 2004, 23:52:25  
 Job time : 196 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2004, 23:38:03 ; Search time 39 Seconds

(without alignments)  
668,583 Million cell updates/sec

Title: US-10-090-185-9

Perfect score: 1388

Sequence: 1 RCLWESRLIQTATAAQAQG.....LNYQLKIKVICIDKSGDVAA 271

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

#### SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1377	99.2	770	2	IGSF3 p91-related
2	1372	98.8	770	2	DNA-binding protei
3	601.5	43.3	739	2	interferon-depende
4	582.5	42.0	748	2	gamma-interferon a
5	364	26.2	851	2	interferon alpha-1
6	288.5	20.8	786	2	mammary gland fact
7	284.5	20.5	793	2	mammary gland fact
8	277.5	20.0	794	2	transcription acti
9	252	18.2	794	2	mammary gland fact
10	129	9.3	837	2	DNA-binding Protei
11	129	9.3	837	2	interleukin-4-indu
12	128	9.2	1208	2	chromosome segrega
13	126	9.1	978	2	conserved hypotnet
14	125.5	9.0	1166	2	hypothetical prote
15	125	9.0	464	2	microtubule bindin
16	124	8.9	533	2	hypothetical prote
17	122.5	8.8	857	2	median body protei
18	122	8.8	1972	1	myosin heavy chain
19	120	8.6	734	2	hypothetical prote
20	120	8.6	1509	1	myosin heavy chain
21	118.5	8.5	924	2	myosin heavy chain
22	118.5	8.5	2094	2	myosin heavy chain
23	118	8.5	284	2	myosin heavy chain
24	118	8.5	2253	2	myosin heavy chain
25	118	8.5	2442	2	myosin heavy chain
26	117.5	8.5	2007	1	myosin heavy chain
27	117	8.4	1818	1	myosin heavy chain
28	117	8.4	1938	1	myosin heavy chain
29	117	8.4	1972	2	myosin heavy chain

30	116.5	8.4	1300	2	153799	CG1 protein - huma
31	116.5	8.4	1356	2	S32763	kinectin 1 - huma
32	116	8.4	289	2	S51193	epimorphin - mouse
33	116	8.4	946	2	S28061	SCP1 protein - rat
34	116	8.4	986	2	T10754	cis-Golgi matrix p
35	115.5	8.3	1178	2	S30431	MSP-300 protein -
36	115	8.3	1164	2	T24806	hypothetical prote
37	114.5	8.2	1690	2	T13030	microtubule bindin
38	114.5	8.2	1999	1	S21801	myosin heavy chain
39	114.5	8.2	3187	2	JC5837	364k Golgi complex
40	114	8.2	1008	2	AE2304	hypothetical prote
41	114	8.2	1410	1	A57013	early endosome ant
42	114	8.2	1961	1	A61231	myosin heavy chain
43	114	8.2	1963	1	MMKW	myosin heavy chain
44	114	8.2	1976	2	A59252	myosin heavy chain
45	114	8.2	2663	1	S28261	centromere protein

#### ALIGNMENTS

##### RESULT 1

IGSF3 p91-related transcription factor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004

C/Accession: I49508; I49009

R/Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su

Cell 77, 63-71, 1994

A/Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr

A/Reference number: A54444; MUID:94208062; PMID:7512451

A/Accession: I49508

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-770 <RES>

A/Cross-references: UNIPROT:P42227; GB:I29278; NID:G476715; PID:AAA37254.1; PID:G47671

R/Raz, R.; Durbin, J.E.; Levy, D.E.

J. Biol. Chem. 269, 24391-24395, 1994

A/Title: Acute phase response factor and additional members of the interferon-stimulate

A/Reference number: I49009; MUID:95014185; PMID:7523373

A/Accession: I49009

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-393, 'M', 395-700, 702-770 <RES2>

A/Cross-references: EMBL:U08378; NID:G473889; PID:AAA56668.1; PID:G473890

C/Genetics:

A/Gene: APRF

C/Superfamily: human signal transducer and transcription activator STAT5A

Query Match 99.2%; Score 1377; DB 2; Length 770;

Best Local Similarity 99.3%; Pred. No. 4.8e-81;

Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	RCLWESRLIQTATAAQAQGQANHPTAAVTEQQMLFQMLQVRRVQDLQKMYE	60
DB	107	RCLWESRLIQTATAAQAQGQANHPTAAVTEQQMLFQMLQVRRVQDLQKMYE	166
QY	61	NLQDDPFRNYKTLSSQGMODLNGNOSVTRQKQOLEQMLTALDQMRSTVSELAGLLS	120
DB	167	NLQDDPFRNYKTLSSQGMODLNGNOSVTRQKQOLEQMLTALDQMRSTVSELAGLLS	226
QY	121	AMEVYQKTLDBELADWKRREPEACIGPPNICDRLQEMWTTSLAESQLQTRQIKXLEE	180
DB	227	AMEVYQKTLDBELADWKRREPEACIGPPNICDRLQEMWTTSLAESQLQTRQIKXLEE	286
QY	181	LQKQVSYGSDPIVQHRPMLERIVLFPNLKSAFVVRQCMRPHRPRVITKGVQFT	240
DB	287	LQKQVSYGSDPIVQHRPMLERIVLFPNLKSAFVVRQCMRPHRPRVITKGVQFT	346
QY	241	TKVRLLVFPELNTQLKIKVICIDKSGDVAA 271	
DB	347	TKVRLLVFPELNTQLKIKVICIDKSGDVAA 377	





A:Accession: A46160  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA; protein  
A:Residues: 1-851 <FUI>  
A:Cross-references: UNIPROT:P52630  
A:Note: sequence extracted from NCBI backbone (NCBI:P110820)  
R.Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.  
submitted to the EMBL Data Library, December 1994  
A:Reference number: S71908  
A:Accession: S71908  
A:Molecule type: DNA  
A:Residues: 1-851 <YAN>  
A:Cross-references: EMBL:U18671; NID:G1293919; P1DN:AAA98760.1; PID:G1293920  
R.Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.  
Nucleic Acids Res. 23, 459-463, 1995  
A:Title: The genomic structure of the STAT genes: multiple exons in coincident sites in  
A:Reference number: S53873; MUID:95192056; PMID:7885841  
A:Accession: S53873  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-196;392-591;684-730 <YAN>  
A:Cross-references: EMBL:U18671  
C:Genetics:  
A:Gene: stat2  
A:Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 40  
C:Superfamily: human signal transducer and transcription activator STAT5A  
C:Keywords: signal transduction; transcription regulation  
Query Match 26.2%; Score 364; DB 2; Length 851;  
Best Local Similarity 33.7%; Pred. No.6.5e-16;  
Matches 89; Conservative 58; Mismatches 107; Indels 10; Gaps 6;  
QY 3 LMESRLIQATATAAQQGQANHPAAVTEKQ-MLEQHLQDVKKRVODLEQKAVEN 61  
DB 112 LLEERKRIIQARQALEQGE--PVLTERPVESQQHEISRLIDRAMWKIVKISQKLD 168  
QY 62 LQDDPDPYKTLKSGQDMQDINGNNQSTRQKQMLEQTLALDOMRSIVSELAGLISA 121  
DB 169 QQDVFCFPRYK-IOAKGKTPSDPH--QTKQKQI-LQETINELDRKREVIDASKALLGR 223  
QY 122 MEVQKTLTDELDLMDKRRPFIACIGPPNICLDRLEWINSLSAQQTQKKEEL 181  
DB 224 LTTLEILL-PKLEWKAQQAQACIRAPIDHGLELFTWAGALLFHLRQLLEKGL 282  
QY 182 QQKVSXKQDPVQHRPMLEERIVLEFRIMKSAFVVERQPCMPHPRPVIKTVGQFTT 241  
DB 283 SCIVSYQDDPLTKGVDLNAQVTELLQRLHRAFAVFERQPCMPHPRPVIKTKGSKRTV 342  
QY 242 KVRLLVKKPELANTQIKTVCTDKD 265  
DB 343 RTRLLVRLQEGNESITVVEISIDRN 366  
RESULT 6  
149274  
mammary gland factor - mouse  
N:Alternate names: STAT5 protein homolog p80  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: 149274; S54773; S54727  
R.Flin, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hemmighausen, L.  
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995  
A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in  
A:Reference number: 149273; MUID:96004632; PMID:7568026  
A:Accession: 149274  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-786 <RES>  
A:Cross-references: UNIPROT:P42232; UNIPROT:Q9JMK1; EMBL:U21110; NID:G747973; P1DN:AA052  
R.Mul, A.L.F.; Makao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.  
EMBO J. 14, 1166-1175, 1995  
A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin  
A:Reference number: S54772; MUID:95237198; PMID:7720707

A:Accession: S54773  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-432, 'E', 434-786 <MUI>  
A:Cross-references: EMBL:Z48539; NID:G758635; P1DN:CA08420.1; PID:G758636  
R.Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Queller, F.; Basu, R.; Saris,  
EMBO J. 14, 1402-1411, 1995  
A:Title: Interleukin-3 signals through multiple isoforms of Stat5.  
A:Reference number: S54725; MUID:95246733; PMID:7537213  
A:Accession: S54727  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-432, 'E', 434-786 <AZA>  
C:Genetics:  
A:Gene: Stat5b  
C:Superfamily: human signal transducer and transcription activator STAT5A  
Query Match 20.8%; Score 288.5; DB 2; Length 786;  
Best Local Similarity 30.0%; Pred. No.4.2e-11;  
Matches 79; Conservative 46; Mismatches 109; Indels 29; Gaps 7;  
QY 1 RC---LMESRLIQATATAAQQGQANHPAAVTEKQMLEQHLQDVKKRVODLEQK 56  
DB 107 RCIRHILYNBQRLVREANNSSPAGS---LADAMSQHLQINQTFEELRLITDTENEL 162  
QY 57 KVENLQDDPDPFNY-KTLKSGQDMQDINGNN-----QSVTRQKQMLEQTL---TAL 104  
DB 163 KKLQQTQEVFIIOQESIRIQARPAQQLQNLPOERMSRFTLMQKQVSELEWLOREAGTL 222  
QY 105 DQMRISVSELAGLISAMEYQKTLTDELDLMDKRRPFIACIGPPNICLDRLEWITS 164  
DB 223 QQYVELEAKHQKTLQLLRKQQTIIILDELLQWRKROQLAANGGPEGSIDLVLQWCERKL 282  
QY 165 AESQLQTRQKKEELQKVSXKQDPVQHRPMLEERIVLEFRIMKSAFVVERQPCMP 224  
DB 283 ARIITWNRQQRRAEHLQOQPIPG-PVEEMLAENVATITDIALVSTFTEKQP--- 338  
QY 225 MHPDRPVIKTVGQFTTKVRLV 247  
DB 339 -----PQVLKQTQTFEATVRLV 356  
RESULT 7  
S54772  
mammary gland factor - mouse  
N:Alternate names: stat5 protein  
C:Species: Mus musculus (house mouse)  
C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S54772; 149273  
R.Mul, A.L.F.; Makao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.  
EMBO J. 14, 1166-1175, 1995  
A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin  
A:Reference number: S54772; MUID:95237198; PMID:7720707  
A:Accession: S54772  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-793 <MUI>  
A:Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:G758633; P1DN:CA08  
R.Flin, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hemmighausen, L.  
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995  
A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved  
A:Reference number: 149273; MUID:96004632; PMID:7568026  
A:Accession: 149273  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-793 <RES>  
A:Cross-references: EMBL:U21103; NID:G747971; P1DN:AAA80590.1; PID:G747972  
C:Genetics:  
A:Gene: Stat5a  
C:Superfamily: human signal transducer and transcription activator STAT5A  
Query Match 20.5%; Score 284.5; DB 2; Length 793;  
Best Local Similarity 29.7%; Pred. No.7.7e-11;

Matches 78; Conservative 46; Mismatches 110; Indels 29; Gaps 7;

QY 1 RC-----LWESRLTQTATATAAQQGQANHPATAVTEKQOMLEQHLQDVKKVQDLQKM 56  
 DB 107 RCIRHILVNEQRLVREANNSSPAG-----LVLDAMSQKHQIQNGFBEHLRLTQDTENEL 162  
 QY 57 KYVENLQDDPDPFN-KTKSGQDMODLGN-----QSVTRQKQQLQOML-----TAL 104  
 DB 163 KKLQQTQEFYFIQYQESLRICQAFQALQQLNPFQEMSEBETALQOKOVSLFTWLQREAO TL 222  
 QY 105 DQMRISIVSELAGLSAMEYVQKTLTDEBLADWKRREPACTIGSPNICTDLLENWITS 164  
 DB 223 QQYRVELAEKHQKTIQLRKQQTITLDBELIQMKRRQQLAGNGPBGSLDVLQSCCKL 282  
 QY 165 AESQLOTRQKIKLELQOKVSYKQDPVIOHRPMLERIVELFRNLMKSAFVERQPCMP 224  
 DB 283 AEIWMNQQRIRRAHILCOQLPIPG-PVEEMLAEVNATITDISLVSTFIIEKQP--- 338  
 QY 225 MHPDRPLVYKIGVQFTTKVRLIV 247  
 DB 339 -----PQVLKTTQTKFAATVRLIV 356

RESULT 8  
 G02317  
 transcription activator stat5a - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
 C:Accession: G02317  
 R:Litn, J.  
 submitted to the EMBL Data Library, December 1995  
 A:Reference number: H01043  
 A:Accession: G02317  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1794 <LIN>  
 A:Cross-references: UNIPROT:P42223; EMBL:U43185; NID:g1151169; PIDN:AA06589.1; PID:g115  
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 20.0%; Score 277.5; DB 2; Length 794;  
 Best Local Similarity 29.3%; Pred. No. 2.2e-10;  
 Matches 77; Conservative 47; Mismatches 110; Indels 29; Gaps 7;

QY 1 RC-----LWESRLTQTATATAAQQGQANHPATAVTEKQOMLEQHLQDVKKVQDLQKM 56  
 DB 107 RCIRHILVNEQRLVREANNSSPAG-----LVLDAMSQKHQIQNGFBEHLRLTQDTENEL 162  
 QY 57 KYVENLQDDPDPFN-KTKSGQDMODL-----NGNNSVTRQKQQLQOML-----TAL 104  
 DB 163 KKLQQTQEFYFIQYQESLRICQAFQALQQLNPFQEMSEBETALQOKOVSLFTWLQREAO TL 222  
 QY 105 DQMRISIVSELAGLSAMEYVQKTLTDEBLADWKRREPACTIGSPNICTDLLENWITS 164  
 DB 223 QQYRVELAEKHQKTIQLRKQQTITLDBELIQMKRRQQLAGNGPBGSLDVLQSCCKL 282  
 QY 165 AESQLOTRQKIKLELQOKVSYKQDPVIOHRPMLERIVELFRNLMKSAFVERQPCMP 224  
 DB 283 AEIWMNQQRIRRAHILCOQLPIPG-PVEEMLAEVNATITDISLVSTFIIEKQP--- 338  
 QY 225 MHPDRPLVYKIGVQFTTKVRLIV 247  
 DB 339 -----PQVLKTTQTKFAATVRLIV 356

RESULT 9  
 S55527  
 mammary gland factor - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
 C:Accession: S55527; S44353  
 R:Nakao, H.; Gouilleux, F.; Groner, B.  
 EMBL J. 14, 854-855, 1995  
 A:Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula

A:Reference number: S55527; MUID:95188889; PMID:7882987  
 A:Accession: S55527  
 A:Molecule type: mRNA  
 A:Residues: 1-794 <WAK>  
 A:Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA5191.1; PID:g60235  
 A>Note: this is a revision to the sequence from reference S44353  
 R:Nakao, H.; Gouilleux, F.; Groner, B.  
 EMBL J. 13, 2182-2191, 1994  
 A:Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcrip  
 A:Reference number: S44353; MUID:94244619; PMID:7514531  
 A:Accession: S44353  
 A:Molecule type: mRNA  
 A:Residues: 1-716; 'RHILGPGSLPSR', 729, 'P', 731, 'ASL', <WAK>  
 A:Cross-references: EMBL:X78428  
 A:Note: this sequence has been revised in reference S55527  
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 18.2%; Score 252; DB 2; Length 794;  
 Best Local Similarity 28.5%; Pred. No. 9.5e-09;  
 Matches 75; Conservative 45; Mismatches 113; Indels 30; Gaps 8;

QY 1 RC-----LWESRLTQTATATAAQQGQANHPATAVTEKQOMLEQHLQDVKKVQDLQKM 56  
 DB 108 RCIRHILVNEQRLVREANNSSAG-----LVLDAMSQKHQIQNGFBEHLRLTQDTENEL 163  
 QY 57 KYVENLQDDPDPFN-KTKSGQDMODLGN-----QSVTRQKQQLQOML-----TAL 104  
 DB 164 KKLQQTQEFYFIQYQESLRICQAFQALQQLNPFQEMSEBETALQOKOVSLFTWLQREAO TL 223  
 QY 105 DQMRISIVSELAGLSAMEYVQKTLTDEBLADWKRREPACTIGSPNICTDLLENWITS 164  
 DB 224 QQYRVELAEKHQKTIQLRKQQTITLDBELIQMKRRHDMRGAEAPR-SLDVLQSCCKL 282  
 QY 165 AESQLOTRQKIKLELQOKVSYKQDPVIOHRPMLERIVELFRNLMKSAFVERQPCMP 224  
 DB 283 AEIWMNQQRIRRAHILCOQLPIPG-PVEEMLAEVNATITDISLVSTFIIEKQP--- 338  
 QY 225 MHPDRPLVYKIGVQFTTKVRLIV 247  
 DB 339 -----PQVLKTTQTKFAATVRLIV 356

RESULT 10  
 I5757  
 DNA-Binding Protein and transcription factor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
 C:Accession: I5757  
 R:Quellle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Cleve  
 Mol. Cell. Biol. 15, 3336-3343, 1995  
 A:Title: Cloning of murine Stat6, Stat proteins that are tyrosine phospho  
 A:Reference number: I5757; MUID:95280934; PMID:7760829  
 A:Accession: I5757  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-837 <RES>  
 A:Cross-references: UNIPROT:P52633; GB:I47650; NID:g1008876; PIDN:AAA79006.1; PID:g100887  
 A:Gene: STAT6  
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 9.3%; Score 129; DB 2; Length 837;  
 Best Local Similarity 25.6%; Pred. No. 0.81;  
 Matches 70; Conservative 33; Mismatches 90; Indels 80; Gaps 13;

QY 14 ATAAQGG-QGA-----NHPTAAVVT-----EQQMLEQ--HLDVKKVQ 50  
 DB 66 ATAGQKGKNGSLPHISTESTYQKDPKLVATITQILQGEKAVIEERHILRGPRHRQ 125  
 QY 51 D-----LEQKRVENIQDDPDPFNKTKSGQDMOD-----LNNNSVTRQKQ 95  
 DB 126 BELKFTPLGRILHVRERTRILRESILGPKT--GVVSLQNLIDPLNPGPS----- 176

QY 96 QLEQMLTALDQMRKRSIVSELGLISAMEYVQKTLTDELDAMKRRPEIACIGPPNICLD 155  
 Db 177 --EDLPITLQ-----GTVGDELTQ-QPLVLRIOIWKQOOLAGNGTFPEBSLA 222  
 QY 156 RLEWMTSLAESQLOTRQOIKLE-ELQOKVSYKGDPIVQHRPMLERIVELFNLKMSA 214  
 Db 223 GLQRCSSLEVIYSQLOHOBISAGSELBPKT-----RASLSRLDELVALTIVTSS 272  
 QY 215 FVEROPCMRHPDRPLVIKTVQFTTKVRLV 247  
 Db 273 FLVEKQP-----POVLKIQTKFQAGVRFLL 297

RESULT 11  
 A54740  
 Interleukin-4-induced transcription factor stat - human  
 C:Species: Homo sapiens (man)  
 C>Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
 C:Accession: A54740  
 R:Hu, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.  
 Science 265, 1701-1706, 1994  
 A:Title: An interleukin-4-induced transcription factor: IL-4 stat.  
 A:Reference number: A54740; MUID:94367369; PMID:8085155  
 A:Accession: A54740  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-848 <HOU>  
 A:Cross-references: UNIPROT:P42226  
 C:Superfamily: human signal transducer and transcription activator STAT5A  
 C:Keywords: DNA binding; transcription regulation

Query Match 9.3%; Score 129; DB 2; Length 848;  
 Best Local Similarity 24.6%; Pred. No. 0.82;  
 Matches 64; Conservative 40; Mismatches 104; Indels 52; Gaps 10;  
 QY 5 EBSRLTQATTAQGGQANHTAAVTEKQMLEHLDVYKRVQDLEQMKVENVLD 64  
 Db 73 EGSYTLQHIHISTL-ESYVRDPLKLVAT-----FRQILQCKKAV-----MEQFRHLP 119  
 QY 65 DFDNRYTKLSQGDMDLNGNQSV--TRQKMOO-----LEQML-----TALDQMR 108  
 Db 120 PFMWQBELKFKTKIRLQHRVGSHILREMLQGAERGVSLSLITPANGTGPSEAL 179  
 QY 109 RSIYSELAGLSAMEYVQKTLTDELDAMKRRPEIACIGPPNICLDRLNMTSLAESQ 168  
 Db 180 AMLIQETTGLELA---AKALVLRIGQWKQOQLAGNAP-----FEESLAPLQ 225  
 QY 169 LQTRQOIKLEELQOKVSYKGDPI-VQHRPMLERIVELFNLKMSAFVERQCMRHP 227  
 Db 226 ERCSLVDIYSQLOQEGVAGSELBPKTASLTGRDELVALTIVTSCFLVEKQP----- 279  
 QY 228 DRPLVIKTVQFTTKVRLV 247  
 Db 280 --POVLKIQTKFQAGVRFLL 297

RESULT 12  
 AE1947  
 chromosome segregation protein [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C:Accession: AE1947  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriguen, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AE1947  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1208 <KUR>  
 A:Cross-references: UNIPROT:O8YXT3; GB:BA000019; PIDD:BAF73085.1; PID:q17130474; GSPDB:G

A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alt1128  
 C:Superfamily: chromosome segregation protein SMCI  
 Query Match 9.2%; Score 128; DB 2; Length 1208;  
 Best Local Similarity 24.0%; Pred. No. 1.4;  
 Matches 62; Conservative 48; Mismatches 94; Indels 54; Gaps 10;

QY 5 EBSRLTQATTAQGGQANHTAAVTEKQMLEHLDVYKRVQDLEQMKVENVLD 50  
 Db 807 QESQLOQLRITLALBLSQTPSEMOCQATIKTQEOQIQRETRFRABQRLKLEHQQ 866  
 QY 51 DLEQMKVENVLDDEDFNYKTLK---SQGDMDLNGNQ--SVTRQKQOQLEQMLTALDQ 106  
 Db 867 RLQRIQEAQQRITVEYQQTTCHEAINRVSSQTTTINAGITQTRAKLSLEHIGAEKQ 926  
 QY 107 MRSIVSELGLISAMEYVQKTLTDELDAMKRRPEIACIG-----PP 150  
 Db 927 KRQITQEGVASHLRQOQLEMEIQKLEETQIKRREDLTALQSQLOELVPELPNLPPEVPD 986  
 QY 151 NICLDRLNMTSLAESQLOTRQOI-----KLEELQOKV-SYKGDPIVQHRP 197  
 Db 987 KVDLELQKELRSIAK-RLQAMEPVNMLAEYERQTKRLBELSQKQTLGE-----RT 1040  
 QY 198 MLEERIVELFNLKMSAF 215  
 Db 1041 ELLRLI-ENFTTLRQIAF 1057

RESULT 13  
 A70387  
 conserved hypothetical protein aq\_1006 - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
 C:Accession: A70387  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
 V.  
 Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70380; MUID:98196666; PMID:9537320  
 A:Accession: A70387  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-978 <AQF>  
 A:Cross-references: UNIPROT:O67124; GB:AB000718; NID:g2983504; PIDD:AAC07092.1; PID:g29  
 A:Experimental source: strain VF5  
 C:Genetics:  
 A:Gene: aq\_1006  
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 9.1%; Score 126; DB 2; Length 978;  
 Best Local Similarity 23.8%; Pred. No. 1.5;  
 Matches 50; Conservative 38; Mismatches 80; Indels 42; Gaps 6;  
 QY 33 EQQQLBQHLQDVYKRVQDLEQMKVENVLDDEDFNYKTLKSGQMDLNGNQSVTRQ 92  
 Db 234 EKKDLSRELSQVVTYTKLEKLEMLEKEVEKLEKLEFSRKVAP-----YVPYAK 281  
 QY 93 KMQQLQMLTALDQMRKRSIVSELGLISAMEYVQKTLT-----DEELAD- 136  
 Db 282 RIEBIDKLTVELKVRKNKLTLEALVLDQESFADEELNRLAEKKEKFEKEREKLEHR 341  
 QY 137 WKRPPEIACIGPPNICLDRLNMTSLAESQLOTRQOIKLEELQOKVSYKGDPIVQHR 196  
 Db 342 LKQLEIKELI-----LKLSSQSSSLKKEKERYEQAKQGFPLSERVE-KGKGLVART 393  
 QY 197 PMLERIVELFRN-----LMKSAFVERQ 220  
 Db 394 EKKLEKIKELFSEEEYTSIMKKEKRLVLELQ 423

RESULT 14

T27075  
 hypothetical protein Y51A2D.16 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #ext\_change 15-Oct-1999  
 C:Accession: T27075  
 R:McMurray, A.  
 Submitted to the EMBL Data Library, January 1998  
 A:Reference number: Z20307  
 A:Accession: T27075  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1166 <MIL>  
 A:Cross-references: EMBL:AL021497; PIDD:CAA16403.1; GSPDB:GN00023; CESP:Y51A2D.16  
 A:Experimental source: clone Y51A2D  
 C:Genetics:  
 A:Gene: CESP:Y51A2D.16  
 A:Map position: 5  
 A:Introns: 17/3; 45/1; 76/2; 156/3; 240/2; 288/2; 594/2; 657/3; 756/3; 798/2; 900/3; 948

Query Match  
 Best Local Similarity 21.4%; Score 125.5; DB 2; Length 1166;  
 Pred. No. 2;  
 Matches 56; Conservative 51; Mismatches 80; Indels 75; Gaps 10;

QY 5 EESRLIQTAATAAQQGGANHTAAVTEKQ-----OMLEQHLD-----VR 46  
 DB 413 DEAKLQPELDEALE--RTSHVTRSLSSSEKNTLKAATLE--LQDOVEAQTLELNOKNGC 468  
 QY 47 KRVODLEQMKVVENLQDFDFNYTKLSQGMODLNGNNSQVTRQKMQOLEQMTALDQ 106  
 DB 469 KRLERDQMSLNLKRLKLENDLTKCQQLLES-----KRLQRLREDLVLEKS 518  
 QY 107 MRSIVSELGLSMEY---VQKLTDEELAD-----WKRPEIACIGGPP 150  
 DB 519 RRADLIGRIHSCTLSLNGANFEKINDDDELIDIMNALVAVYKREDDRLIQGNQ 578  
 QY 151 NI-----CIDRLNNITSLAESQLOTRQOI-----KCLELQOKVSY 187  
 DB 579 QIQEHLDKRLDEKLRSESESLNSDDRVRELTRNNMTTKQVFMLOEKLELNLLEIST 638  
 QY 188 KGDPIVQHRPMLERIVELFRN 209  
 DB 639 KNDEI-----DMVKASIEELNRN 656

RESULT 15  
 H90279  
 microtubule binding protein, probable [imported] - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
 C:Accession: H90279  
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.  
 Submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: J99139  
 A:Accession: H90279  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-464 <KUR>  
 A:Cross-references: UNIPROT:Q9UXN4; GB:AB006641; NID:g13814451; PIDD:AAK41495.1; GSPDB:C  
 C:Genetics:  
 A:Gene: S801256

Query Match  
 Best Local Similarity 9.0%; Score 125; DB 2; Length 464;  
 Pred. No. 0.72;  
 Matches 73; Conservative 69; Mismatches 107; Indels 88; Gaps 15;

QY 3 LWSERLLQTAATAAQQGGANHTAAVTEKQMLEQHLDV-----KKRVODLEQK 55  
 DB 92 LEESTKLEQAVQELIEAQKHDERITKLESTKLEQAVQELIEAQKHDERITKLEES 151  
 QY 56 MKVVEN-LQDFDFNYK-----TLKSQGMQDL-----NGNNSQVTR----- 91

DB 152 TKLEQAVQELIEAQKHDERITKLEESTKLEQAVQELIEAQKHDERITKLESTKXL 211  
 QY 92 -QKMOOL-----ROMLTALDQMRSTYSSELGLSMEYVQKTLTD-----EELADPK 138  
 DB 212 EQAVQELIEAQKHDERITKLEESTKLEQAVQELIEAQKHDERITKLEESTQKLVDAQ 271  
 QY 139 RRPEIACIGGPPNICIDRLNNITSLAESQLOTRQOIKLEELQKV--SYKG--DPIVQ 194  
 DB 272 RRAEER-----LAKLENAVEQLVEAQKTRITKLEEVTKLVESQLVSGQMNRE 322  
 QY 195 HRPML-----EERIVELFRNLK-----SAFVVERQPCMPHEDDPLVTK 234  
 DB 323 LRKALGSMGKRWGRDPFEKLIIEIVDELAKQEGLDLKYKKFTYK-----DNGLFGL 374  
 QY 235 TGVQFTTKVRLVKKPELNYOLKIKVCIDKDSGVAA 271  
 DB 375 KQVEY--DVDLILKDTKV-YLIRIKSYVEKDVMAA 408

Search completed: December 2, 2004, 23:53:10  
 Job time : 42 secs